Heriot-Watt University
Research Gateway

Scan-to-BIM for ‘secondary’ building components

Citation for published version:
Adán, A, Quintana, B, Prieto, SA & Bosché, FN 2018, 'Scan-to-BIM for ‘secondary’ building components', *Advanced Engineering Informatics*, vol. 37, pp. 119–138. https://doi.org/10.1016/j.aei.2018.05.001

Digital Object Identifier (DOI):
10.1016/j.aei.2018.05.001

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Peer reviewed version

Published In:
Advanced Engineering Informatics

Publisher Rights Statement:
© 2018 Elsevier B.V.

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

Scan-to-BIM for ‘secondary’ building components

Antonio Adán[1], Blanca Quintana[1], Samuel A. Prieto[1] and Frédéric Bosché[2]

[1] 3D Visual Computing and Robotics Lab
Universidad de Castilla-La Mancha.
Ciudad Real, Spain
{Antonio.Adan, Blanca.Quintana, Samuel.Prieto}@uclm.es

[2] Institute for Sustainable Building Design, CyberBuild Lab
Heriot-Watt University
Edinburgh, U.K.
f.n.bosche@hw.ac.uk

Abstract: Works dealing with Scan-to-BIM have, to date, principally focused on 'structural' components such as floors, ceilings and walls (with doors and windows). But the control of new facilities and the production of their corresponding as-is BIM models requires the identification and inspection of numerous other building components and objects, e.g. MEP components, such as plugs, switches, ducts, and signs. In this paper, we present a new 6D-based (XYZ + RGB) approach that processes dense coloured 3D points provided by terrestrial laser scanners in order to recognize the aforementioned smaller objects that are commonly located on walls. This paper focuses on the recognition of objects such as sockets, switches, signs, extinguishers and others. After segmenting the point clouds corresponding to the walls of a building, a set of candidate objects are detected independently in the colour and geometric spaces, and an original consensus procedure integrates both results in order to infer recognition. Finally, the recognized object is positioned and inserted in the as-is semantically-rich 3D model, or BIM model. The assessment of the method has been carried out in simulated scenarios under virtual scanning providing high recognition rates and precise positioning results. Experimental tests in real indoors using our MoPAD (Mobile Platform for Autonomous Digitization) platform have also yielded promising results.

Keywords: object recognition, scan-to-BIM, automatic BIM, 3D data processing.

1. RELATED WORK.

Building Information Modelling (BIM) is now rapidly penetrating the Architecture, Construction, Engineering and Facilities Management (AEC&FM) sectors. However, it is notable that this increase in use has been predominantly for new builds, as well as on the design stage of those new constructions [1], [2] . Yet, it has long been argued that the most significant value of BIM will be delivered during the FM stage [3]. Enabling this is not just a matter of convincing clients and FM teams of the value of BIM information for their Repair and Maintenance (R&M) activities, but also ensuring that the information they receive is indeed comprehensive and accurate [4], [5]. In the case of new builds, this means that Asset Information Models (derived from BIM models) passed on to clients are the result of as-built/as-is asset state analysis (as opposed to a blind distribution of the as-designed BIM model).

In the case of existing builds, this means that a comprehensive and accurate as-is AIM model must be produced that accurately capture the asset as-is. Focusing on the scope of information contained in the BIM model, it is particularly valuable to R&M teams that their AIM contains not just architectural information, but also structural as well as Mechanical, Electrical and Plumbing (MEP) information. Indeed, MEP costs typically constitute the largest share of R&M costs. Therefore, any AIM should not just contain objects like walls with openings (doors and windows), but also many other building service components, such as sockets, safety lights or alarm devices, etc. In the remainder of this manuscript, we shall refer to these objects as secondary building services' components (SBSC).

The efforts made as regards automatic Scan-to-BIM modelling from laser scanning or photogrammetric data have, to date, been focused on segmenting and recognizing large architectural 'structural' components such as walls (with openings), floors, ceilings and columns. For the reasons stated above, it is clearly necessary to extend automatic Scan-to-BIM beyond these large 'primary' components to other smaller 'secondary' ones, such as secondary building service components.

The automatic detection and localization of secondary building service components in point cloud data is a challenging research line that has, however, been of comparatively little interest to the research community to date. Apart from the works that detect openings (door, windows and moldings) in walls using 2D and/or 3D information [6]–[9], only a few works deal with the recognition of other smaller, building service components, such as luminaries, sockets and others. Moreover, the detection is frequently associated with recognizing and positioning objects in a particular 2D space (i.e. an image of the scene) that is not connected with any as-is semantically-rich model (i.e. BIM model) of the building.

With regard to mechanical equipment, most works focus on pipe recognition and positioning. For example, Czerniawski et al. ([10]) recognize a specific pipe spool type in cluttered point cloud scans. A 3D CAD of the search query, and more specifically, its local curvature characterization, is set from the beginning of the search. The points with a similar curvature are filtered from the original point cloud and then clustered by using a density based clustering algorithm [11]. A bag-of-features matching procedure finds the best hypothesis and selects the recognized group. The underlying assumption of the method is that the query pipe spool curvature pattern must be very different to that of the surrounding clutter.

Focusing now on electrical equipment, several researches take advantage (assume) of the fact that lights, luminaries and other heat-emitting resources can easily be detected in thermal images. Kim et al. [12] propose a method that detects in thermal 3D point clouds hot and cool regions on ceilings, but crudely assume that these regions correspond to electrical systems, heating, ventilation and air-conditioning (HVAC) components; no recognition/classification algorithm is employed. In [13], a thermal point cloud is first segmented by temperature, with each cluster representing an object class. The objects in a cluster are then classified by means of a decision tree classifier based on geometric features. Three obviously distinct classes (humans, screens and light fixtures) are clustered in this work. Since a high temperature gradient appears around the query objects in all cases, the segmentation stage is quite effective.

Other authors distinguish fluorescent lighting tubes from circular low energy bulbs by processing 2D coloured images of ceilings [14]. In a first step, a binary image of the ceiling is obtained by mean of a rasterization algorithm of the corresponding 3D data points. In a second step, and after applying Harris's corner detector and the Hough transformation to the image, the authors discriminate several zones of the ceiling, which are assumed to be either fluorescent lighting tubes or circular low energy bulbs lights. But, like in [12], no proper recognition/classification algorithm is employed, which raises questions a regards the generalisability of the method. In [15], another luminary detection approach is presented for

application in tunnels. Several restrictions are imposed in this proposal. For instance, it is assumed that the luminaries are located at a certain height and separated from each other by specific distances. After selecting a 3D data slice at the imposed height, the lights are easily identified by applying a simple colour thresholding technique.

With regard to wall-mounted electrical equipment, like sockets and switches, [16] presents the detection of two types of electrical outlets in coloured images acquired by a robot, which are then classified as power holes, ground holes and the background. Since it is assumed that the regions around the holes do not have any texture, the authors apply an intensity-based method for recognition. The 3D coordinates are detected by applying a planar PnP solver. Meeussen et al. ([17]) propose a method that also employs mobile robots to recognize doors, door handles, electrical plugs and sockets in an office environment. Door handles are recognized by means of the separate processing of digital images, range data and point clouds. Outlet detection is carried out using disparity images, point clouds and colour images.

In [18], sockets and switches are recognized in orthographic 2D images. Object detection is performed with a simple sliding window and patch matching approach. The probability of detection is measured using a feature descriptor pool (a kind of HoG that models the distribution of image gradients in different ways) and a random forest classifier is used to classify power sockets, light switches and the background. Kang et al. ([19]) detect lift call buttons using a stereo camera. First, an adaptive thresholding generates a binary image of the lift. Some buttons and the floor number candidates are then found in this image. Ambiguous candidates are rejected using an artificial neural network, and finally, a matching method is applied to recognize not only the call buttons, but also more properties, such as destination floor buttons, the direction in which the lift is moving and its current location. This is obviously a previous stage that must subsequently be solved within a more complex robot manipulation problem.

Another interesting work is that of Hamledari et al. [20]. The proposed algorithm detects four kinds of objects in 2D images of walls under construction: insulation, studs, electrical outlets and different states for drywall sheets. This information could provide valuable information for progress tracking systems. However, since the method is applied to 2D images, the recognition results are not integrated into a 3D (BIM) model, which leads us to believe that its applicability is currently restricted to the mere monitoring and visualization of these objects on walls.

The work presented by Bonanni et al.[21] is a human-robot collaboration approach that is designed to extract 3D shapes associated with objects of interest. In this case, the objective of the method is to recognize three different objects: fire extinguishers, hydrant boxes and printers. In this work, the system requires human intervention. After the scanner has acquired a 6D (XYZ-RGB) image of the scene, the human uses a laser pointer to point out the region in which the query object lies. The segmentation process then extracts the data that belong the object.

2. CONTRIBUTIONS OF THE METHOD

Our approach is original in that it has been designed to generate semantically-rich 3D models of buildings that include important secondary building services' components. More specifically, when compared to the aforementioned approaches, the main contributions of our work are:

I. It is an automatic method which neither necessitates human interaction nor has demanding requirements. With some methods, the user has to set the class of object to be recognized ([14], [21]), while others fix particular geometric properties of the query objects ([15], [22]).

II. All of the aforementioned methods detect very few objects. Only one or two different objects are recognized in ([12], [14]–[16], [21], [22]) and three or four in [13],[23], [18]. Our

approach is able to detect a larger number of objects because it is based on a learning recognition algorithm. In the experimental section, we consider thirteen different objects.

III. Our object recognition method uses separate algorithms based on colour and on depth, and then establishes a consensus rule between both results. This makes the method more robust and flexible when dealing with different 3D shapes and textures. This is an original idea, which has never been seen in the earlier works.

IV. Our approach is integrated within a larger Scan-to-BIM system so far focused on 'primary' architectural/structural elements, so that the recognised secondary objects are accurately inserted and positioned into the 3D semantically-rich model of the building (i.e. BIM model). This is an important contribution to the automatic Scan-to-BIM research line. As the literature review showed, no work has yet been reported that integrates so many secondary building service components into a Scan-to-BIM solution.

The paper is organized as follows. Section 3 briefly explains the prior data processing stages, including data acquisition, the segmentation of visible areas on walls and the detection of openings. The core of the paper comprises the object recognition approach. This is explained throughout the four sub-sections of Section 4. Recognition with geometry and recognition with colour are presented separately in sub-sections 4.2 and 4.3. The consensus decision algorithm is subsequently addressed in Section 4.4. Section 5 presents the experimental work, which includes the assessment of the technique in both a simulated building and a real environment. Our conclusions and future improvements are summarized in Section 6.

3. DATA ACQUISITION, PRE- PROCESSING/PREPARATION

3.1. Previous steps: data acquisition and segmentation

The input of our object recognition system is a coloured point cloud associated with an already detected and modelled wall. In our case, we get this information from our existing Scan-to-BIM solution focused on primary architectural/structural components [24].

In the data acquisition stage of our solution, a mobile mapping system for digitization (MoPAD) takes data from the environment. Our autonomous moving scanner collects sufficient information to roughly represent the interior of the building. At this stage, the 3D model consists of an unstructured set of points that represents the visible scene.

Data segmentation and labelling of the essential constructive elements of the building follows. This is a step in which a semantic meaning is introduced into the data. We generate a point cloud model composed of segments in which the objects "wall", "ceiling", "floor" and "column" have a semantic meaning. The extraction of the points belonging to the floor and ceiling of the room is carried out first. This is easily done by detecting two maximums in the Z-histogram of the data. We assume here that ceilings and floors are planar and parallel regions.

The segmentation of the points belonging to each wall of an indoor scene is conducted afterwards. The point cloud is first projected onto the XY plane and is then discretized, thus generating a 2D image *I*. After finding the edges of the polygon that encloses the data in *I* and returning to the 3D space, the 3D data related to the walls are retrieved. All details of the automatic scanning and Scan-to-BIM approach for primary architectural/structural components can be found in [24].

3.2. Data preparation

The set of coloured data belonging to each constructive element (i.e. a wall) is structured as a 4D orthoimage, J_{CD}, in which each pixel has colour (RGB) and depth (i.e. an orthonormal

distance between the 3D points and the wall plane). The resolution of J_{CD} is 5mm × 5mm per pixel, signifying that the image resolution is low. Note that this could be the standard resolution for 3D laser scanners used in the indoor scanning of buildings.

After registering several views (coloured scans) and, since different natural and artificial light sources may impact on a wall from different locations in the room, slight colour variations (with boundaries) usually appear in the visible wall areas. Additional, and very important, specular highlights resulting from the system's camera flash could greatly distort the colour in some zones of the image. Under these circumstances, simple object detection algorithms based on colour thresholding would probably not be effective. In order to reduce such colour variations and restore highlights zones of the raw orthoimage, we have pre-processed the 4D orthoimage. We refer the reader to [25] for a better understanding of this stage.

3.3.Detection of the wall area and openings

In order to detect secondary building components in walls, we first detect the visible areas of the walls and the existing openings (windows and doors). The algorithm that detects visible areas on walls finds clusters of coherent colour seeds on the wall and then carries out a segmentation by colour. As a result of this process, the visible area of the wall is separated from the rest of the wall and the openings on it are afterwards sought. The approach that recognizes doors and windows has recently been published in [25]. A brief explanation is provided here.

The detection of openings is based on finding discontinuities in the 4D RGB-D space. We process the colour and depth components of J_{CD} separately, with J_{CD} being decomposed into J_C (colour) and J_D (depth), after which the results are recombined.

Given our assumption that the door frames are rectangular, we detect straight lines in J_{CD}. These lines represent the discontinuities as regards the colour and depth of the wall (if the door has a protruding doorframe, the discontinuity in the depth dimension should result in line detections; if the door is a different colour from that of the wall, the discontinuity in the colour dimensions should also result in line detections). Note that the lines detected contain parts of the contours of hypothetical doors owing to potential occlusions.

All possible rectangles defined by two pairs of horizontal and vertical lines are found. Since we are seeking rectangles that delimit openings, we retain only those rectangles whose size falls within the range of typical opening sizes. This yields a highly reduced set of rectangles. Finally, each rectangle is recognized as an actual opening if it fulfils a set of conditions regarding properties of colour and depth consistency, the degree of door frame occlusion and location consistency within the wall. Figure 1 d) shows the result obtained after applying the method. Note that this is not a simple example because of the multiple glass regions in a double door and the signs fixed on it.

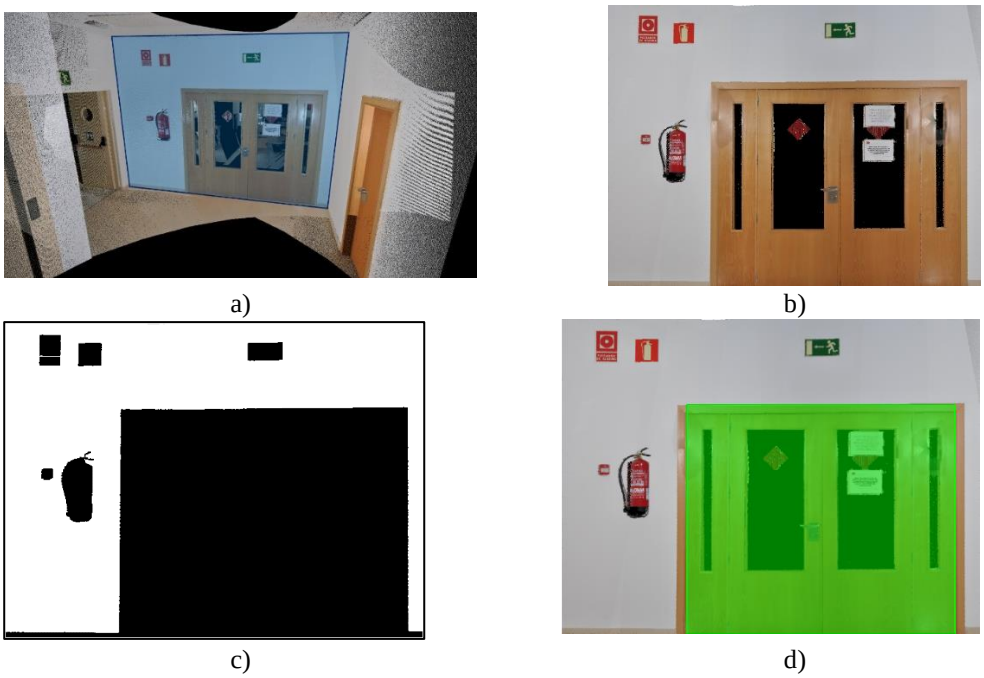


Figure 1.a) Total registered coloured point cloud and the data segment of a particular wall (highlighted in blue). b) Resulting orthoimage, J_{CD}. c) Detection of the visible wall area in white. d) Recognition of the door in the wall.

4. RECOGNITION OF BUILDING SERVICE COMPONENTS: THE APPROACH

4.1. An overview

Figure 2 a) shows a general flowchart with the main stages explained in Sections 2 and 3, whereas Figure 2 b) provides a detailed outline of the object recognition method proposed in this paper. Our approach assumes the existence of a database of objects that may be present in the query building. This database includes colour and depth image models for each object. Note that, in the case of a Scan-vs-BIM scenario, this database could be automatically generated using the as-design BIM model of the facility. The as-design BIM model would also provide additional valuable information, such as which objects should be present on each wall. In this paper, we very much place ourselves in such a context.

The flowchart shown in Figure 2 starts splitting the aforementioned orthoimage J_{CD} into image J_C (colour) and J_D (depth), which will be further processed separately. Upon removing the existing openings from images J_C and J_D, they become new images $\hat{J}_C$ and $\hat{J}_D$. Potential regions of interest (RoI) are subsequently calculated in both images, respectively.

Different object recognition strategies are then applied to the corresponding RoIs. In both cases, a model database is used to obtain a tentative list for each query object class in the scene. After applying the recognition algorithm, some classes will be in both lists (share recognition), others will be in only one list (exclusive recognition) and the rest will not be in either list (non-recognized).

This is followed by the consensus stage. Bearing in mind that there could be several instances of the same class in the scene and several candidates per class for each strategy, a *Recognition Coherence Matrix* Ψ is computed for each class. Each entry of Ψ is the so-called *Recognition Coherence Level*, which is the measure of the coherence between detecting the object class in images $\hat{J}_C$ and $\hat{J}_D$ at their respective calculated positions. The recognition consensus decision will be solved by choosing the best values of matrix Ψ, and a precise

position of each recognized object instance on the wall will be calculated by means of a weighted mean position formula. All of the above will be explained in Section 4.4.

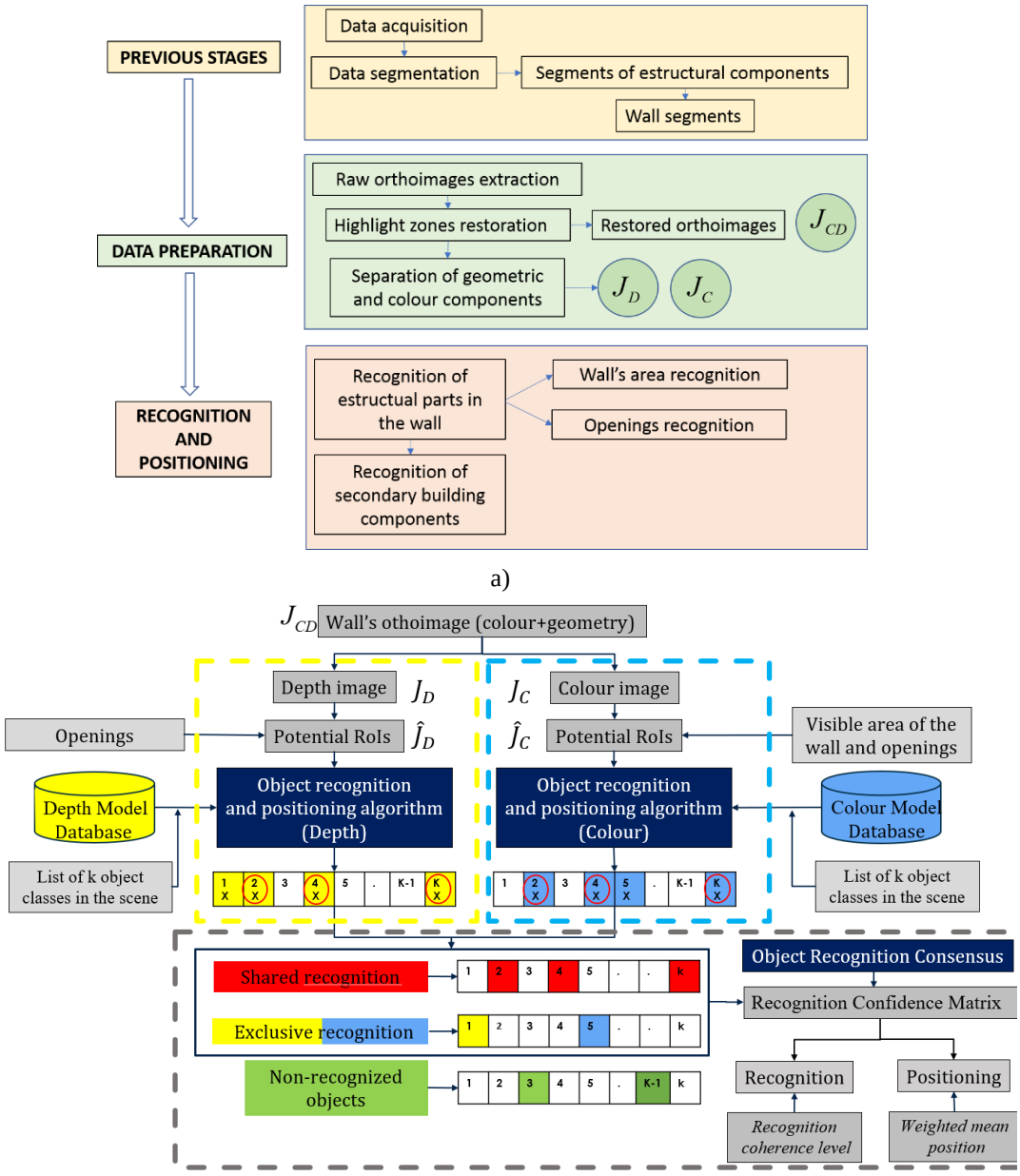


Figure 2. a) General overview with previous stages, data preparation and recognition. b) Flowchart resuming the object recognition approach.

4.2. Recognition with geometry

As mentioned in the previous section, after detecting and positioning the existing openings in the orthoimage J_{CD} , we generate a new 4D orthoimage $\hat{J}_{CD}$, which contains only the information contained in the wall area, and this is decomposed into a depth image $\hat{J}_D$ and a colour image $\hat{J}_C$.

$\hat{J}_D$ is used to detect objects with geometric discontinuities with regard to the wall plane, which is valuable for the detection of protruding objects. The recognition approach follows two stages. Firstly, $\hat{J}_D$ is used to calculate potential RoIs (rectangles) containing geometric discontinuities. Secondly, a matching stage between the RoIs and the depth models of the database is carried out. Figure 3 provides an example in which only one pair of objects has depth discontinuities.

Calculation of RoIs

In order to find salient regions on the wall, a Canny filter is applied to $\hat{J}_D$ (Figure 3 b)). The Canny algorithm computes by itself two thresholds to detect strong and weak edges. By using two thresholds, the Canny method is less sensitive to noise than other edge detection techniques and more effective to detect true weak edges. The resulting image $\hat{J}_{D,Canny}$ is processed with the objective of detecting closed boundary regions, which are later enclosed by rectangles, each of which represents a RoI. The RoIs of the orthoimage in Figure 1 are shown in Figure 3 c).

As a result of this process, objects with a complex geometry might be bounded by several overlapping rectangles, each of which corresponds to a different part of the object. A clustering algorithm integrates different overlapping RoIs into a single RoI. The initial set of candidate RoIs is thus reduced (see Figure 3 e)).

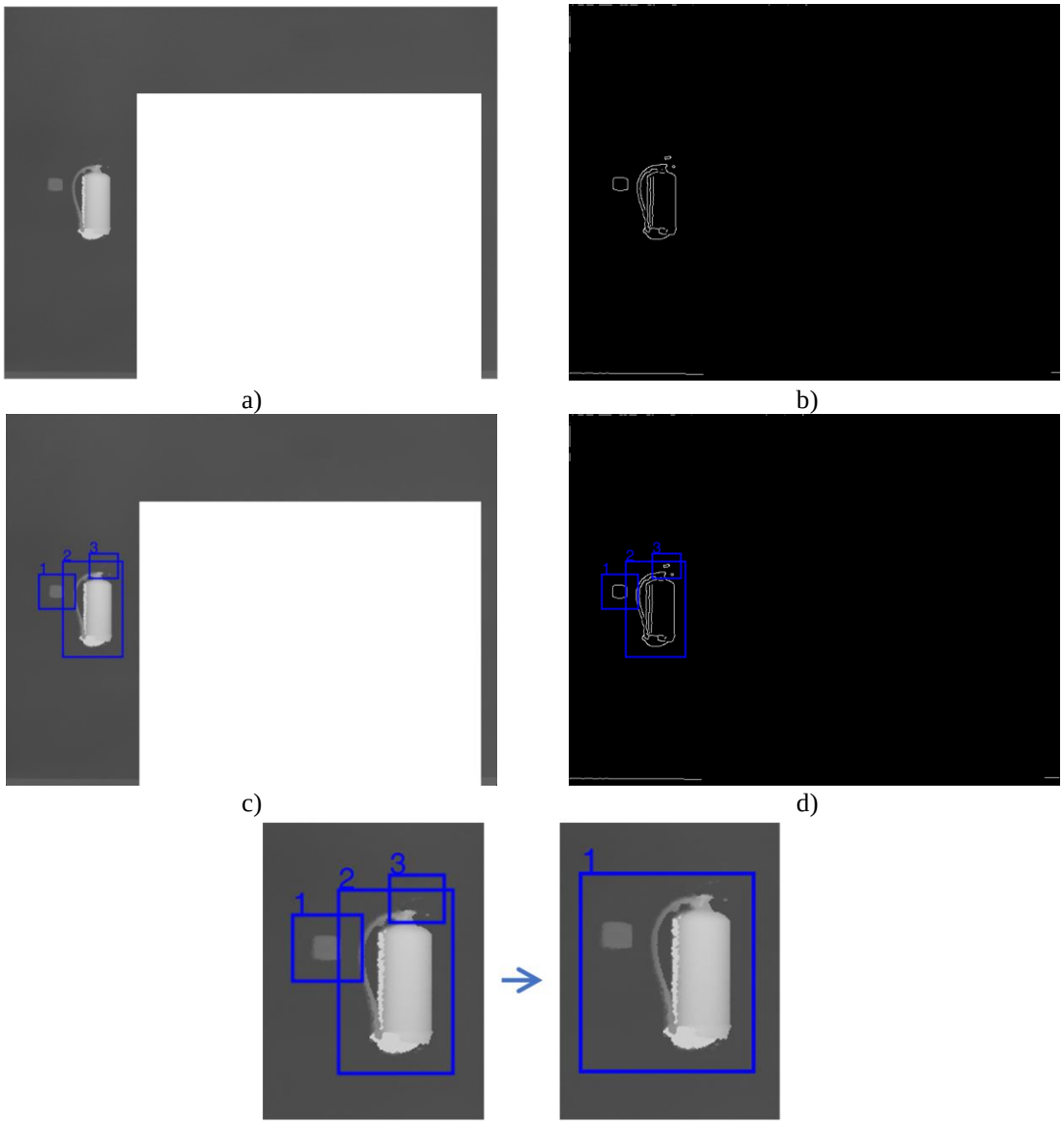


Figure 3. a) Image $\hat{J}_D$. b) Image $\hat{J}_{D,Canny}$. c) Calculated RoIs superimposed onto $\hat{J}_D$. d) Calculated RoIs superimposed onto $\hat{J}_{D,Canny}$. e) RoIs reduction by clustering overlapping bounding boxes.

Matching Stage

In the second stage, an image cross-correlation algorithm assesses the correlation between the images contained in the RoIs and the depth models of the database. Note that we assume the identity of the objects in the scene in advance, and these will from here on be denominated as "query or expected objects". We correlate each expected object model with the set of calculated RoIs and obtain a list of matched RoIs, all overtaking a cross-correlation coefficient threshold μ_D, ($\mu_D = 0.75$). Thus, for an expected object O_i, we eventually extract a list of RoIs in the depth-space,$\{D\}_{O_i}$

An example of this process is shown in Figure 4. Note that the input 4D orthoimage (and, therefore, $\hat{J}_D$) and the depth models are both implemented on the same scale in pixel/cm, signifying that it is not necessary to conduct any cross-correlation operations on multiple scales. Note that the cross-correlation function also yields the precise position of the best matching in $\hat{J}_D$. The blue rectangles in Figure 4 b) and c) identify potential RoIs that would match an extinguisher and a fire alarm switch. Each step of the object recognition algorithm is presented in detail in Algorithm I.

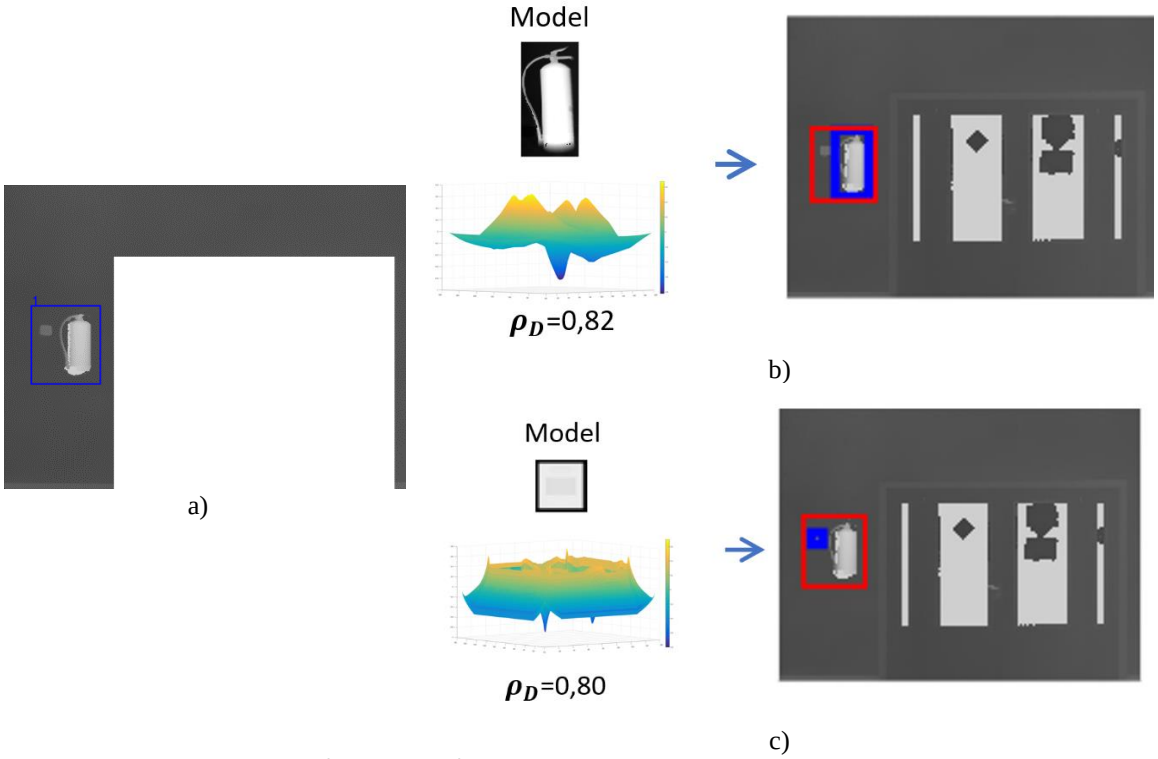


Figure 4. Recognition results in $\hat{J}_D$. a) RoI to $\hat{J}_D$. b) and c) Detail of the cross-correlation map of extinguisher 2 and fire alarm switch. The blue rectangles identify the objects on the wall.

```
Input: Orthoimage J_CD, set of doors and windows detected and Object
Models
Output: List of RoIs in the depth-space {D}_Oi

begin //
generation of orthoimage Ĵ_CD
decomposition of Ĵ_CD into  Ĵ_C and Ĵ_D
// calculation of RoIs
Ĵ_D,Canny ← (Canny filter, Ĵ_D)
Candidate Regions ← (closed boundaries, Ĵ_D,Canny)
RoIs of Candidate Regions
RoIs clustering
// matching stage
for each Object Model do
       for each RoI do
              ρ^Model_D,Candidate ← Cross Correlation algorithm
              if ρ^Model_D,Candidate > μ_D
                     Matched RoI ← Object Model
                     List of RoIs {D}_Oi ← Matched RoI
              elseif
                     NOT Matched RoI ← Object Model
              end if
       end for
end for
end //
```

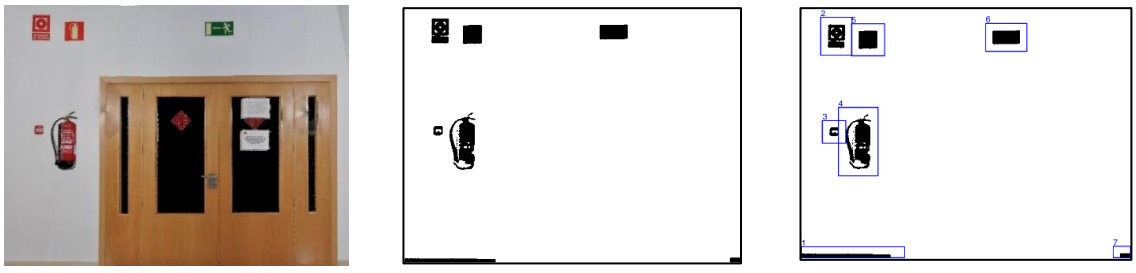
Algorithm I. Object recognition algorithm using geometry.

4.3. Recognition with colour

Image $\hat{J}_C$ is used to detect objects as colour discontinuities in the wall area, and objects that are salient in the colour domain are, therefore, easily detected. The first stage again calculates RoIs, in this case by means of discontinuities in the colour domain. A matching algorithm between the colour models and the calculated RoIs is then carried out.

Calculation of RoIs

We first discard the visible wall area from $\hat{J}_C$, so that it contains only those parts of the wall that are neither the visible wall area nor openings. The resulting image, $\hat{J'}_C$, is subsequently transformed into a binary image $\hat{J}_{BW}$ (as illustrated in Figure 5 b)) that is processed with the aim of finding compact sets of pixels (black regions), each of which theoretically represents a candidate region to be dealt with. $\hat{J}_{BW}$ is obtained by taking a threshold defined from the mean intensity level of the previously extracted wall area. As in the case of the recognition depth based approach, the RoI (bounding rectangle) is calculated for each candidate region and stored. Figure 5 c) shows image $\hat{J}_{BW}$ and the RoIs calculated.



(a) (b) (c)

Figure 5. (a) Image J_C, (b) image $\hat{J}_{BW}$ and (c) RoIs calculated in the segmentation process.

Definition of the pattern

The RoIs are matched into the model database by means of a minimum distance-based classifier using global descriptors. Although local descriptors are commonly used for object

recognition in coloured images (SIFT, SURF, among others), these techniques provide frustrating results when applied to our orthoimages of 5mm/pixel. Note that we have to balance the resolution of the collected point cloud with the associated memory and time requirements. Although the theoretical angular stepwidth between consecutive scan lines ranges from 0.0024º to 0.5º in our Riegl VZ-400, an intermediate value provides sufficient resolution, in a reasonable time, as when extracting precise 3D point models of buildings. Bearing all of the above in mind, we fixed the angular stepwidth at 0.065º, which yields 10 million points in 83 seconds per 360-scans. However, owing to this limited resolution, along with the colour distortions produced after the registration of several coloured point clouds from different scanner locations, small objects within our orthoimages appear blurred and are of a poor quality. This could be a frequent problem in the point cloud processing world.

Figure 6 illustrates the poor quality and low resolution of different objects in the orthoimage corresponding to Figure 1. Note that, for example, the image of the fire alarm switch extracted from J_C is 26x26 pixels in size.

For the aforementioned reasons, and with the objective of developing an effective object recognition technique for such poor quality coloured images, we have defined a set of global descriptors which are invariant to scale and rotation. A training set is used to learn each pattern prototype.

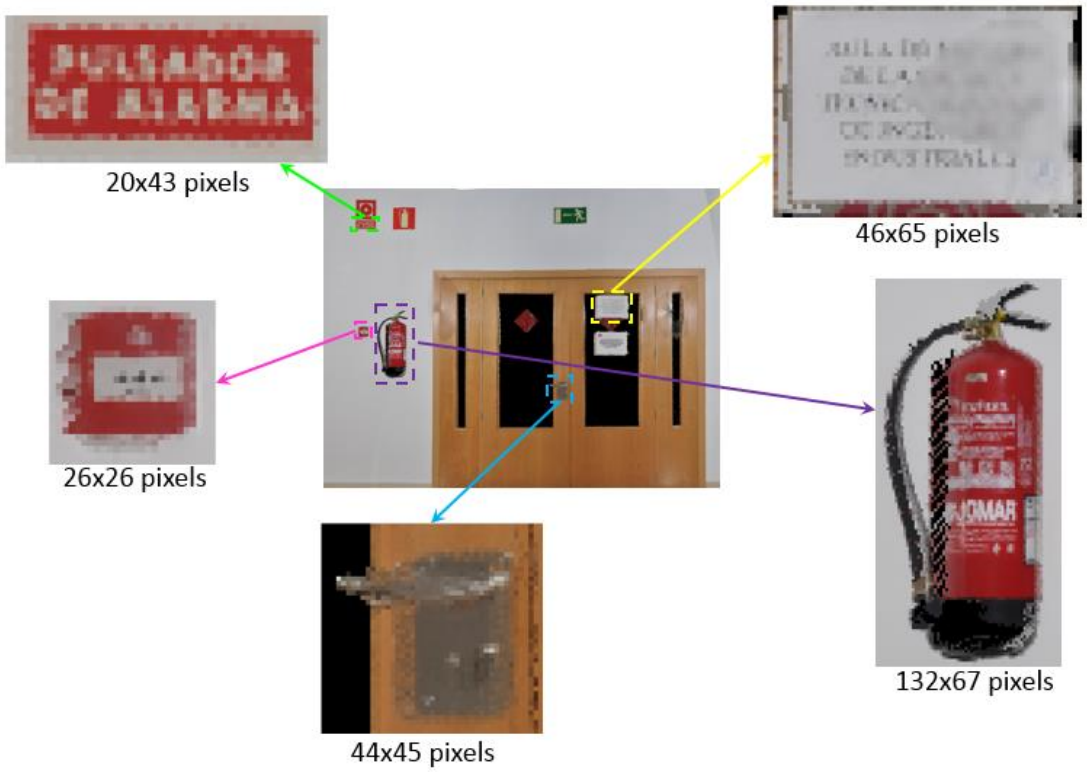


Figure 6. Details of image J_C , which illustrate the low resolution of different objects on the wall.

In order to train the recognition system, we take five front views per object and calculate a pattern $V = (v_1, v_2, \ldots v_{15})$ composed of 15 global descriptors, most of which are related to the HSV and Lab colour spaces. We carried out extensive experimentation and evaluated the use of many other global descriptors related to the colour and geometric shapes of objects in blurred scenes. We eventually concluded that the best results appear with simplified colour palettes, relative colour relationships and global shape descriptors.

After separating the foreground and background in the object image, we ran a drastic clustering algorithm for the components "saturation" (HSV), "*a"* (Lab) and "*b"* (Lab). We

found that saturation remains constant under rotation and scale in images and is not significantly sensitive to slight illumination changes. As is known, in the LAB space, *a* represents (roughly) redness versus greenness and *b* represents yellowness versus blueness. Since they are both relative colour measures, the use of *a* and *b* is also less sensitive to errors and blurred images. *S*, *a* and *b* were, therefore, used to define the first thirteen descriptors as follows.

Components *S*, *a* and *b* of the original image are clustered into three classes and the first and second class prototypes are used to define four global characteristic. Let I_{3S}, I_{3a} and I_{3b} be the corresponding images, each composed of three classes (see Figure 7 d)) and $\{s_1, s_2\}$, $\{a_1, a_2\}$ and $\{b_1, b_2\}$ be the respective first and second class prototypes in the respective images. The first five descriptors of the pattern *V* are defined with the saturation prototypes $\{s_1, s_2\}$ as follows:

$$v_1 = s_1 \tag{1}$$

$$v_2 = s_2 \tag{2}$$

$$v_3 = {\langle s_1 \rangle}/{\langle I_{3S} \rangle}, \text{ where the symbol } \langle\,.\,\rangle \text{ signifies ordinal} \tag{3}$$

$$v_4 = {\langle s_2(s_1) \rangle}/{\langle s_1 \rangle} \tag{4}$$

A brief explanation follows. v_1. And v_2 are the two principal saturation values in I_{3S}. v_3 is the percentage of pixels s_1 in I_{3H} (not considering those corresponding to the background). v_4 is the percentage of pixels s_2 contained in regions s_1. We have additionally included descriptor v_5 as the number of segments s_2 contained in regions s_1.

Figure 7 illustrates all the steps, from the original RGB image to the images of the first and second *S*-prototypes.

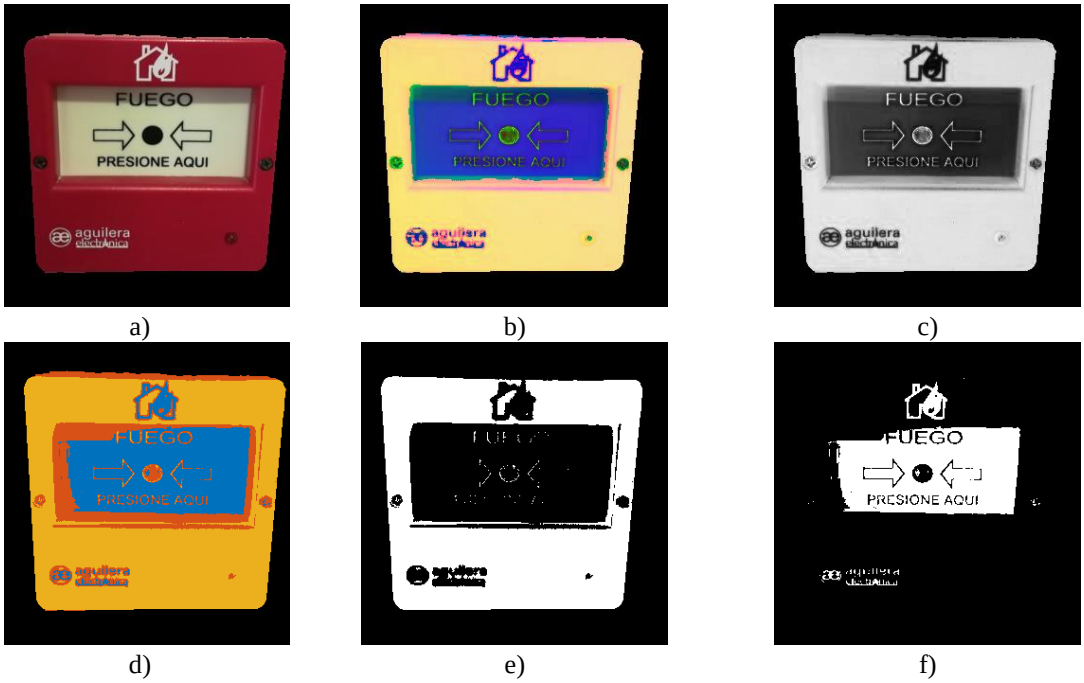


Figure 7. a) RGB sample image, I_{RGB}. b) Image in the HSV space, I_{HSV}. c) Saturation component. Image I_S. d) Saturation reduced to three clusters. Image I_{3S}. e) Image showing the first S-prototypes$_1$ in white. f) Image showing the second *S*-prototypes$_2$ in white.

Descriptors v_6 to v_9 and v_{10} to v_{13} are defined in the same way as the earlier v_1 to v_4, but now for prototypes $\{a_1, a_2\}$ and $\{b_1, b_2\}$ in the Lab colour space.

$$v_6 = a_1 \tag{5}$$

$$v_7 = a_2 \tag{6}$$

$$v_8 = \langle a_1 \rangle / \langle I_{3a} \rangle \tag{7}$$

$$v_9 = \langle a_2(a_1) \rangle / \langle a_1 \rangle \tag{8}$$

$$v_{10} = b_1 \tag{9}$$

$$v_{11} = b_2 \tag{10}$$

$$v_{12} = \langle b_1 \rangle / \langle I_{3b} \rangle \tag{11}$$

$$v_{13} = \langle b_2(b_1) \rangle / \langle b_1 \rangle \tag{12}$$

Figure 8 and Figure 9 show the earlier images in the Lab space and their corresponding a and b-prototypes.

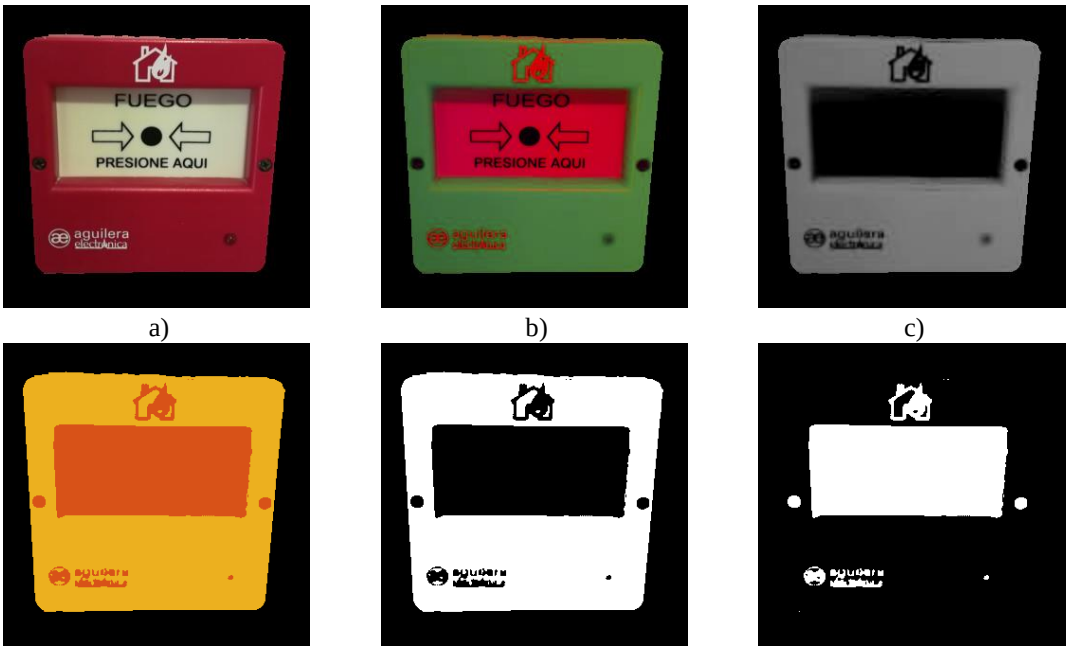


Figure 8 a) RGB sample image, I_{RGB}. b) Image in the Lab space, I_{Lab} . c) a-component. Image I_a . d) Image reduced to three clusters. Image I_{3a}. e) Image showing the first *a*-prototype a_1 in white. f) Image showing the second *a*-prototype a_2 in white.

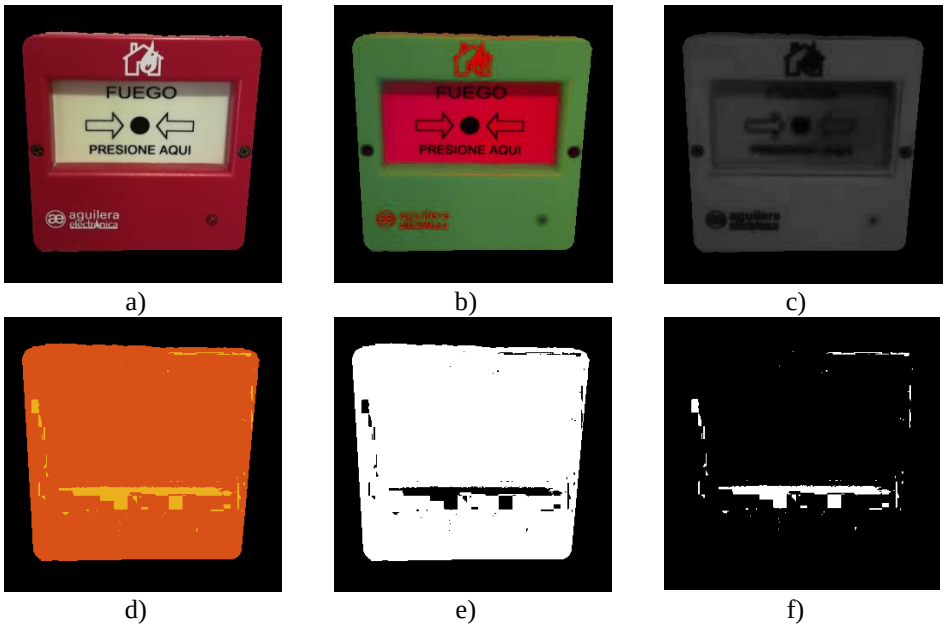


Figure 9. a) RGB sample image, I_{RGB}. b) Image in the Lab space, I_{Lab} . c) b-component. Image I_b . d) Image reduced to three clusters. Image I_{3b}. e) Image showing the first *b*-prototype b_1 in white. f) Image showing the second *b*-prototype b_2 in white.

Finally, the shape descriptors are the roundness and the energy formulated in the equations below.

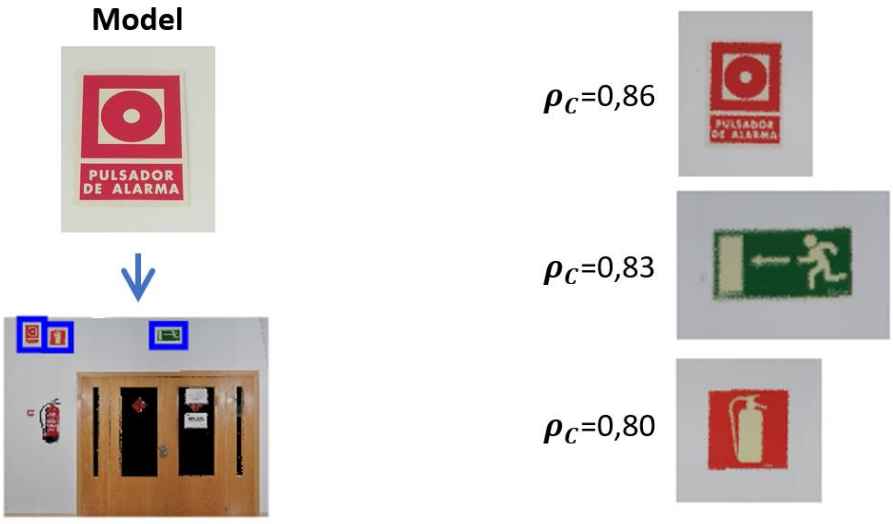
$$v_{14} = \frac{4\pi A}{P^2} \tag{13}$$

$$v_{15} = \sum_{i,j} p(i,j)^2 \tag{14}$$

v_{14} is the roundness of the object, where A is area and P is perimeter, and v_{15} is the energy descriptor. This is calculated as the sum of the squared elements ($p(i,j)$ in Equation 14) in the grey-level co-occurrence matrix (GLCM).

Applying a Minimum Distance Classifier

As in the case of recognition with geometry, we match each query object pattern with the set of patterns corresponding to the previously extracted RoIs by using a minimum distance classifier (Euclidean distance). Finally, a list of matched RoIs, all with distances below a threshold μ_C, ($\mu_C = 0.2$) is obtained. Thus, for a query object O_i, we eventually extract a list of RoIs $\{C\}_{O_i}$. An example of the RoIs matched to a particular sign model is presented in Figure 10. Algorithm II presents the steps of the algorithm in detail.



a) b)

Figure 10. a) Example of matched RoIs obtained for a model after applying the minimum distance classifier. b) Distances $\boldsymbol{\rho_c}$ of the RoIs candidates calculated.

Input: Orthoimage J_{CD}, set of doors and windows detected and Object Models
Output: List of RoIs in the depth-space $\{C\}_{O_i}$

begin //
generation of orthoimage $\hat{J}_{CD}$
decomposition of $\hat{J}_{CD}$ into $\hat{J}_C$ and $\hat{J}_D$
// calculation of RoIs
$\hat{J}'_C \leftarrow$ *removing previously detected wall area from* $\hat{J}_C$
$\hat{J}_{BW} \leftarrow$ binarization of $\hat{J}'_C$
Candidate Regions ← processing of $\hat{J}_{BW}$ finding compact objects
RoIs of Candidate Regions ← bounding boxes of Candidate Regions
// matching stage
System training ← Database model samples
for each RoI **do**
    Background removing
    Descriptors Calculation
    // Minimum distance algorithm
    Calculate distance ρ_C^{Model} to the models ← Descriptors
    if $\rho_C^{Model} < \mu_C$
        Matched RoI ← Object Model
        List of RoIs $\{C\}_{O_i}$ ← Matched RoI
    elseif
        NOT Matched RoI ← Object Model
    end if
end for
end //

Algorithm II. Object recognition algorithm using colour.

4.4. Consensus Strategy for Recognition and Positioning

Although some objects might be detected by means of both geometry and colour, some others will be recognized only by using colour or geometric properties. For example, signs hung on walls are probably detectable only in the colour space, and white electrical switches on white walls might be detectable only with geometric properties. However, extinguishers are probably recognized in both geometry and colour data. Furthermore, several instances of the same object might be found on the same wall.

As was explained in sub-sections 4.2 and 4.3, for each expected object, two lists of candidates $\{D\}_{O_i}$ and $\{C\}_{O_i}$are obtained from both recognition algorithms. In order to find a consensus between the results of depth-based and colour-based recognitions, in a context of multiple instances, the following method is proposed.

From $\{D\}_{O_i}$ and $\{C\}_{O_i}$, a *Recognition Coherence Matrix* Ψ is calculated (see Figure 11). Each entry in $\Psi(O_i)$ is the *Recognition Coherence Level* α, which measures the coherence between a pair of RoIs in $\hat{J}_D$ (or none) and $\hat{J}_C$ (or none). For example, $\alpha(D1, C3)$=0.65 signifies that the recognition of object O_i from the RoIs *D1* in $\hat{J}_D$ and *C3* in $\hat{J}_C$ has a coherence level of 0.65.

α	C1	C2	C3	...	Cn	No color
D1	0.0	0.76	0.65	...	0.74	0.5
D2	0.65	0.98	0.6	...	0.0	0.5
...	...	...	...	...	...	...
Dn	0.0	0.0	0.0	...	0.0	0.5
No Depth	0.5	0.5	0.5	...	0.5	0

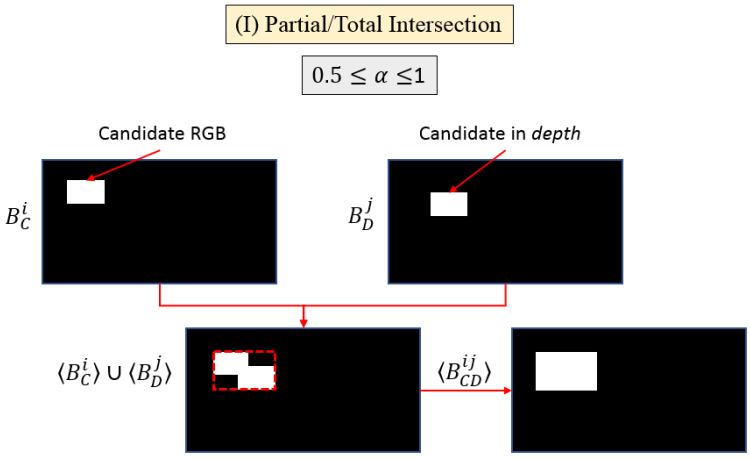
Figure 11. Example of Recognition Coherence Matrix **Ψ** for a certain object.

The *Recognition Coherence Level* between two candidates is calculated by assessing the overlap between the two RoIs, B_C^i and B_D^j (with centroid coordinates c_C^i and c_D^j respectively), as:

$$\alpha = \begin{cases} (I) & \dfrac{\langle B_C^i \rangle \cup \langle B_D^j \rangle}{\langle B_{CD}^{ij} \rangle} \quad if\ B_C^i \cap B_D^j \neq \emptyset \\ (II) & 0 \quad if\ B_C^i \cap B_D^j = \emptyset \\ (III) & 0.5 \quad if\ \nexists B_C^i or \nexists B_D^j \end{cases} \tag{15}$$

where B_{CD}^{ij} is the bounding box that encloses B_C^i andB_D^j, and $\langle B \rangle$ is the number of pixels inside the bounding box B. Note that $\alpha \in [0,1]$. For a better understanding, Figure 12 provides an example of B_C^i, B_D^j, $\langle B_C^i \rangle \cup \langle B_D^j \rangle$, $\langle B_{CD}^{ij} \rangle$ and the corresponding value of α for cases (I), (II) and (III) in Equation 15.

If the RoIs are overlapping, α takes high values (next to 1), otherwise α progressively decreases as the distance grows (case I). If the RoIs are not overlapping $\alpha = 0$ (case II). In those cases in which the object is detected in only one of the two images, we set α as 0.5 (case III).



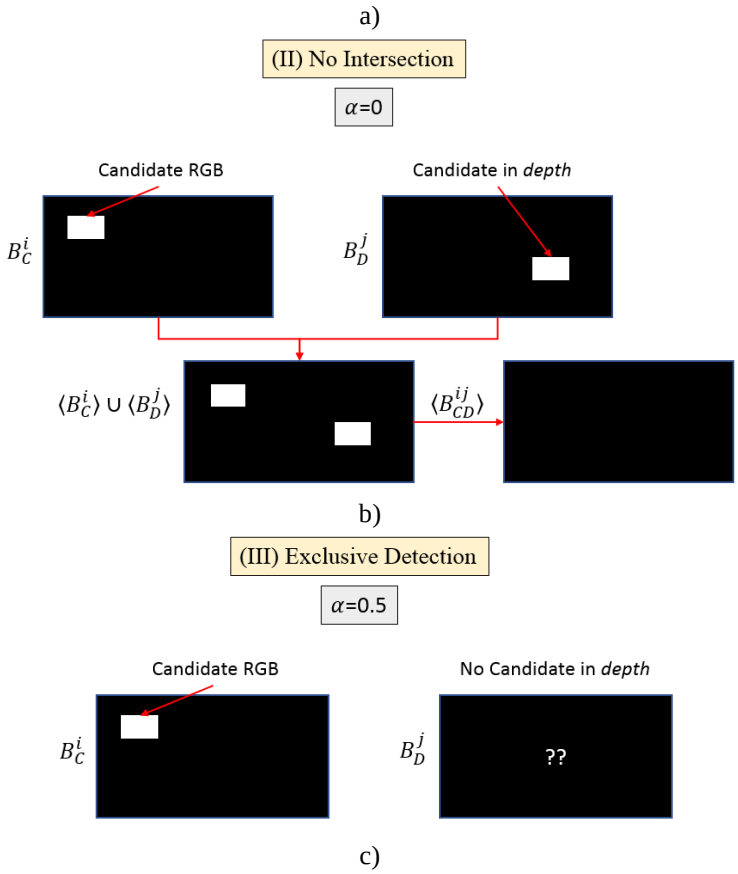


Figure 12. Illustration of $\boldsymbol{B_C^i}$, $\boldsymbol{B_D^j}$, $\langle \boldsymbol{B_C^i} \rangle \cup \langle \boldsymbol{B_D^j} \rangle$, $\langle \boldsymbol{B_{CD}^{ij}} \rangle$ and the corresponding value of α for cases a) (I) partial/total intersection, b) (II) no intersection and c) (III) exclusive detection.

Once Ψ has been filled, the recognition consensus decision is solved iteratively as follows (see Figure 13 for a better understanding):

(1) the highest value of Ψ is selected and considered to be a recognized instance of the query object.

(2) the corresponding row and column of Ψ are eliminated, except when the selected cell corresponds to an exclusive detection case (case III), in which only the corresponding cell is set to 0.

(3) the process is iterated until Ψ is null or until the number of selected cells is equal to the number of expected instances of the query object.

The position (centre) of each instance of the recognized object in the orthoimage J_{CD}, c, is then calculated using Equation 16:

$$c = \frac{\rho_D c_D + \rho_c c_C}{\rho_D + \rho_C} \tag{16}$$

Where c_C and c_D are the centroids of B_C^i, B_D^j and the two weights ρ_D and ρ_C are the cross correlation coefficient (which evaluates the goodness of the recognition in the depth image) and the minimum distance coefficient (which evaluates the goodness of the recognition in the colour image). In the case of exclusive detection, one of the values ρ_D and ρ_C will be zero in Equation 16.

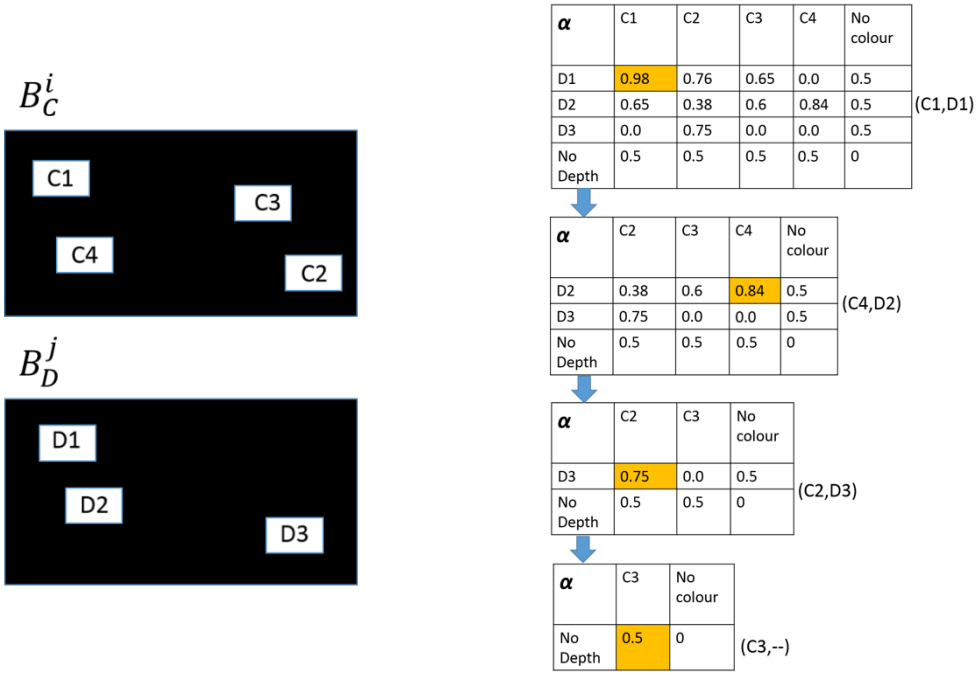


Figure 13. a) Example of multiple instances of an object in images in $\hat{J}_C$ and $\hat{J}_D$. b) The *Recognition Coherence Matrix*, Ψ and the recognition consensus process. After four iterations, four instances of the object are recognized in positions calculated from Equation 16 for the RoI pairs (C1,D1), (C4,D2) and (C2,D3) and in C3.

5. EXPERIMENTAL RESULTS

In this section, we present the experimental results obtained after using our secondary building component recognition approach. The experimentation was carried out in two different datasets in simulated and real scenarios.

5.1.Simulated Scenarios

Scenario I.

The simulated scenario is shown in Figure 14. It consists of the floor of a building of 22.9m x 19.4 m in size, composed of 4 rooms and a corridor. The scenario contains 116 secondary building objects located on 24 walls, which have been painted in neutral colours. This maintains a moderate colour contrast between each wall and their corresponding SBSCs.

The objects are small common objects in buildings, such as extinguishers, signs, switches, sockets or radiators, among others. Some details of different walls with small objects can be seen in Figure 14 c). The coloured point cloud was obtained using Blensor [26], a tool integrated into the Blender modeller that is able to simulate 3D scanners. This software tool allows us to carry out the automatic scanning of the scene modelled, thus enabling us to obtain coloured point clouds similar to those obtained with our Riegl VZ-400 3D laser scanner. Blensor can also add noise to the position and colour of the point cloud collected, which makes the experimentation realistic. The object model database has also been generated with Blensor.

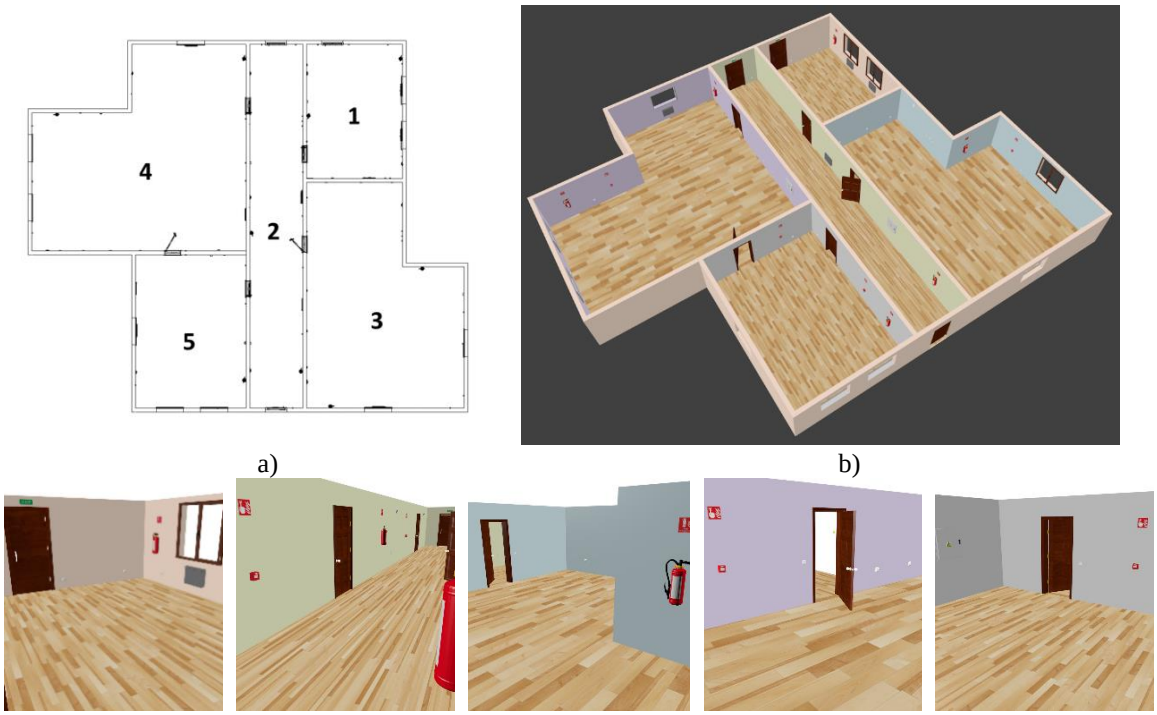
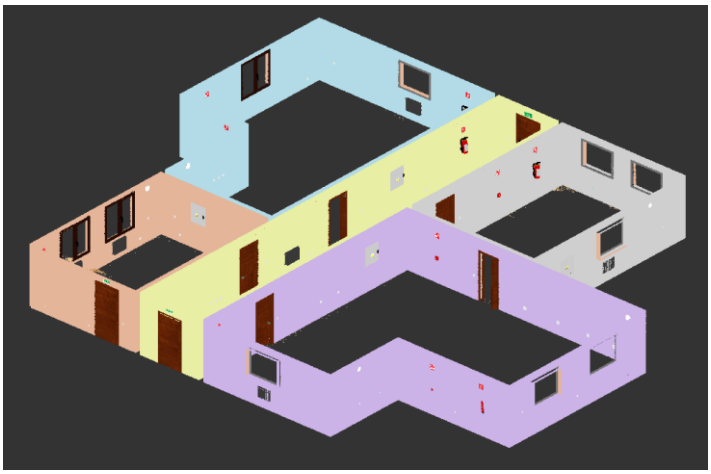
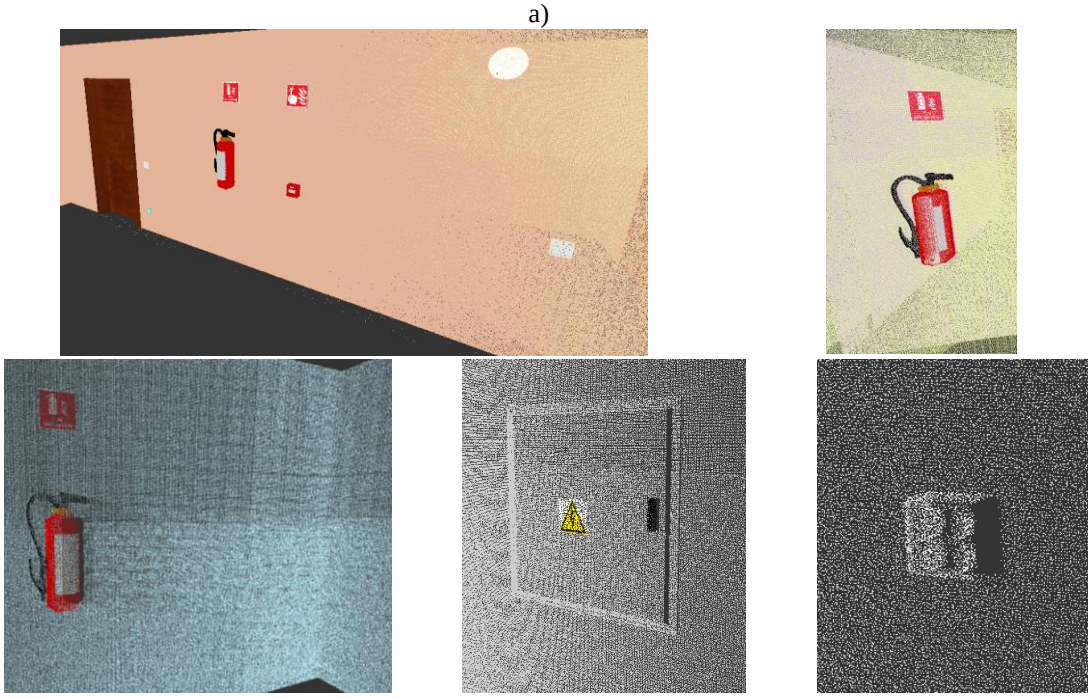


Figure 14. Scenario I. a) Blueprint of the building floor. b) 3D view of the scenario. c) Details of some walls with small building objects.

Figure 15 shows the point cloud obtained with Blensor. The points on the ceiling and floor have been removed for a better visualization. In Figure 16, some examples of the orthoimages (colour and depth) generated from the collected point cloud are illustrated. Note the poor quality and slight discontinuities in the depth images for some objects, particularly, signs, switches and sockets. They will hardly be identified in the depth images and will only be recognizable in colour images. Furthermore, some objects of a colour that is similar to that of the wall, will be recognized only in depth images (see Figure 16 d)). In Figure 16 b) and c) shows some areas with a lack of data (black areas in colour images and white areas in depth images). This lack of data entails discontinuities in the images and might lead to errors in the RoI identification stages.

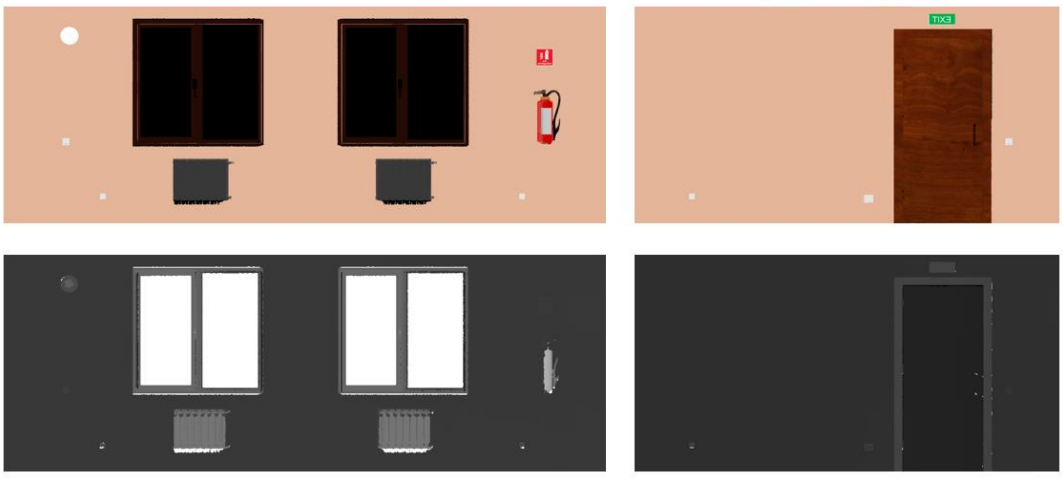


a)

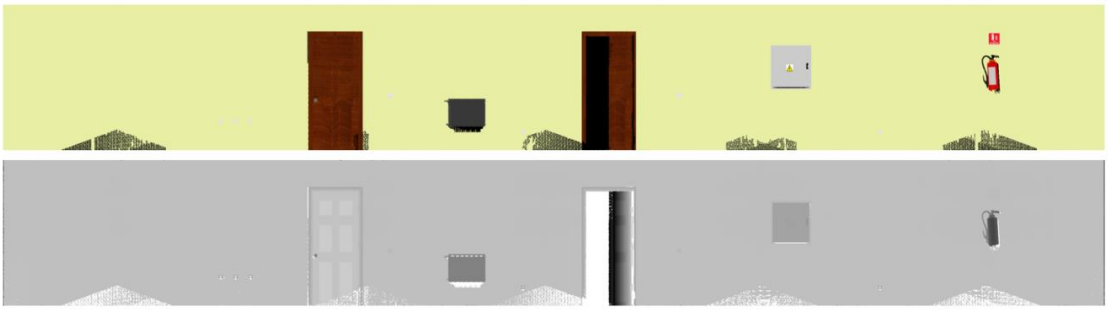


b)

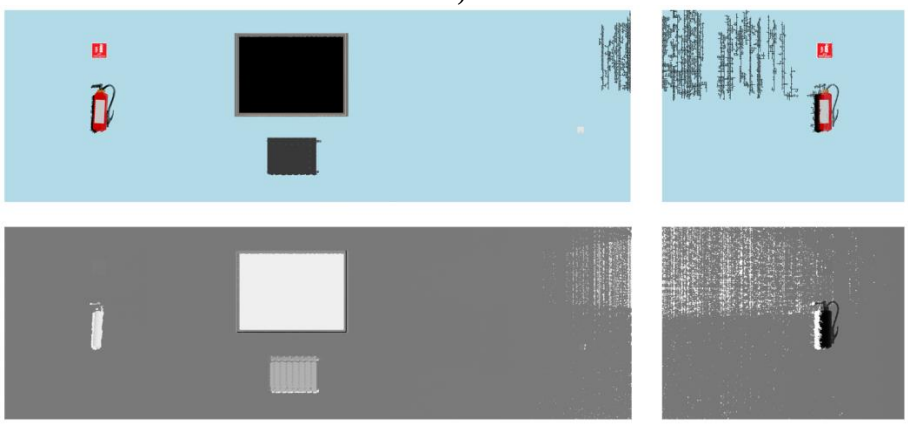
Figure 15. a) Coloured point cloud model of the simulated scenario. Floor and ceiling have been removed for a better visualization. b) Details of the coloured point cloud.



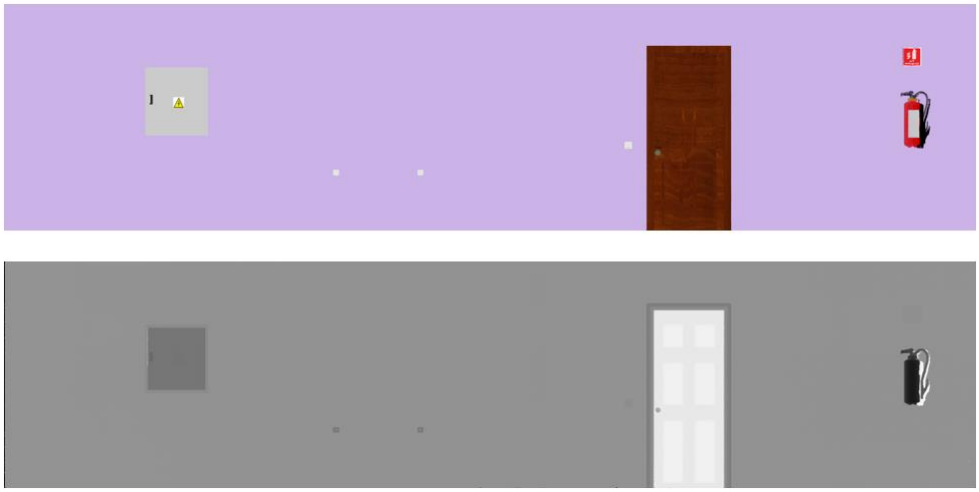
a)



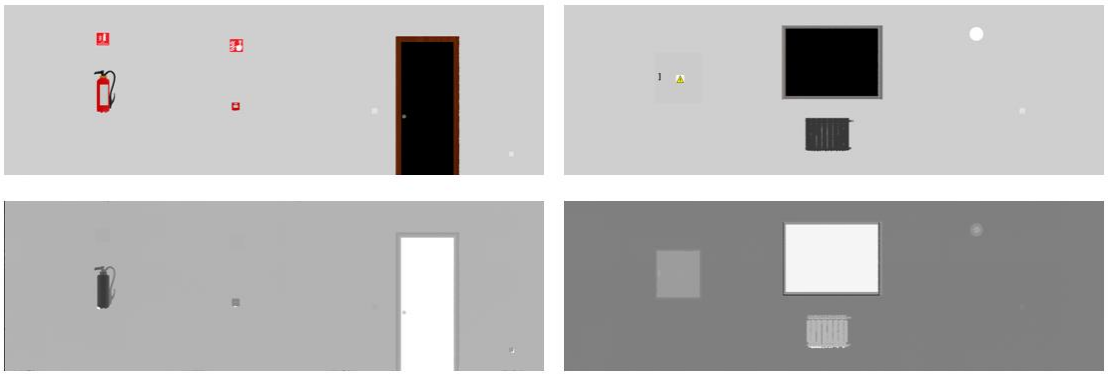
b)



c)



d)



e)

Figure 16. Examples of orthoimages (colour and depth) of a) Room #1, b) Room #2, c) Room #3, d) Room #4 and e) Room #5.

Table 1 shows the results obtained for the simulated scenario. Our approach recognized 105 out of 116 objects, signifying that the percentage of true positives was 90.5%. The wrong results were: 10 false positives (8.6%) and 11 false negatives (9.5%). We also measured the accuracy of the calculated object positions with regard to the ground truth. The horizontal and vertical mean errors between the calculated and ground truth centroids were 2.9mm and 2.7mm, respectively. Table 1 shows more details of the values *Tp*, *Fn* and *Fp* per object class.

The sixth column contains the percentages for which the true positives have a *Recognition Coherence Level* α that is higher than 0.9. Note that low values correspond to objects that are recognized either by colour or by geometry (e.g. extinguisher sign, fire alarm switch sign, radiator), whereas high values are those identified in both orthoimages, $\hat{J}_D$ and $\hat{J}_C$, (e.g. socket x2, switch). The seventh and the eighth columns show the mean horizontal and vertical position errors per object class. The last two columns correspond to the maximum horizontal and vertical errors. It is noteworthy that the maximum horizontal and vertical errors are just 30mm and 27mm, respectively. These results demonstrate the accuracy and the good performance of our approach.

Figure 17 shows a graph of the *Recognition Coherence Level* α for each recognized object. The true positives are represented in blue, and the false positives in red. Note that α is 0.5 for all the false positives, with the exception of one.

Table 1. Scenario I. Recognition results for building service components

Object	Number of instances	TP (%)	FN (%)	FP (%)	TP (%) α>0.9	$\overline{\Delta_h}$ (mm)	$\overline{\Delta_v}$ (mm)	Max Δ_h (mm)	Max Δ_v (mm)
Electrical Panel	4	75.0	25.0	0.0	66.7	5.0	1.7	6	3
Socket x1	20	95.0	5.0	5.0	31.6	2.1	1.7	5	5
Socket x2	6	100.0	0.0	0.0	100.0	2.0	3.0	4	5
Socket x4	11	100.0	0.0	0.0	63.6	2.3	2.3	4	3
Built-in Socket	6	83.3	16.7	16.7	100.0	1.2	1.4	3	3
Switch	16	100.0	0.0	0.0	81.3	2.1	1.8	5	4
Fire Extinguisher	9	100.0	0.0	0.0	33.3	9.9	5.9	30	10
Radiator	6	100.0	0.0	0.0	16.7	3.0	2.2	8	6
Fire Alarm Switch	8	62.5	37.5	37.5	60.0	1.8	2.0	6	3
Smoke Detector	10	80.0	20.0	20.0	87.5	3.1	1.4	10	3
Exit Light	3	100.0	0.0	0.0	0.0	1.3	16.0	2	27
Extinguisher Sign	9	88.9	11.1	11.1	0.0	2.8	2.0	8	5
Fire Alarm Switch Sign	8	75.0	25.0	25.0	0.0	2.0	1.5	4	4
TOTAL/Average	116	90.5	9.5	8.6	50.5	2.9	2.7	30	27

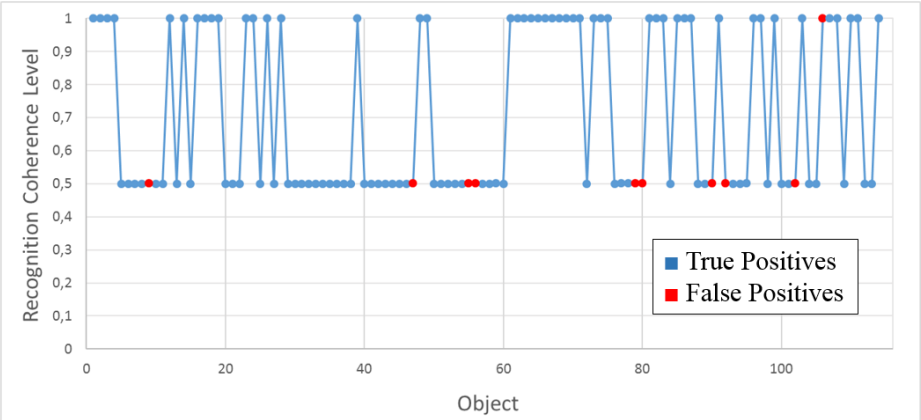
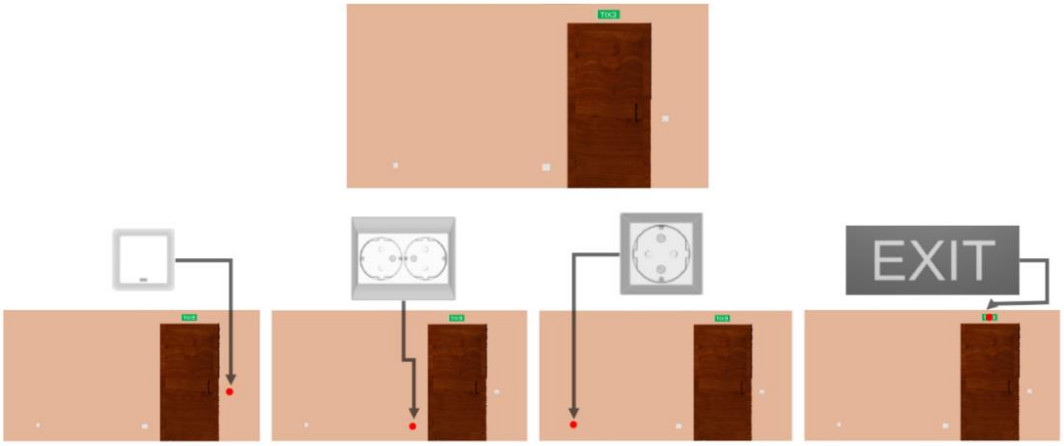


Figure 17. Graph of the *Recognition Coherence Level* α for each recognized object.

Figure 18 shows examples of recognized objects for three representative walls. The red spots indicate the position of the recognized objects. All the objects in rooms #1 and #3 have been correctly recognized. In Room #2, the electrical panel has not been recognized. This failure is due to the fact that the cross correlation coefficient ρ_D (0.69) is less than the imposed threshold μ_D (0.75). On the other hand, the colour-based recognition algorithm was not able to find the corresponding RoI candidate region. Note that even though some objects (e.g. the socket in Room #2) are near non-sensed areas (areas with a lack of data), which produces erroneous discontinuities, they are eventually correctly recognized.

Figure 19 shows the results for the complete simulated scenario. Figure 19 a) illustrates the 3D position of the recognized objects (red spots) together with the structural elements, doors and windows previously recognized. In Figure 19 b), the previous image is superimposed onto the simulated scenario in Blender, in such a way that it is possible to compare the ground truth and the results obtained with our approach. The spots corresponding to the different objects are represented in different colours for each room. Some details of this image are shown in Figure 19 c) for a better visualization.

Room 1



Room 2

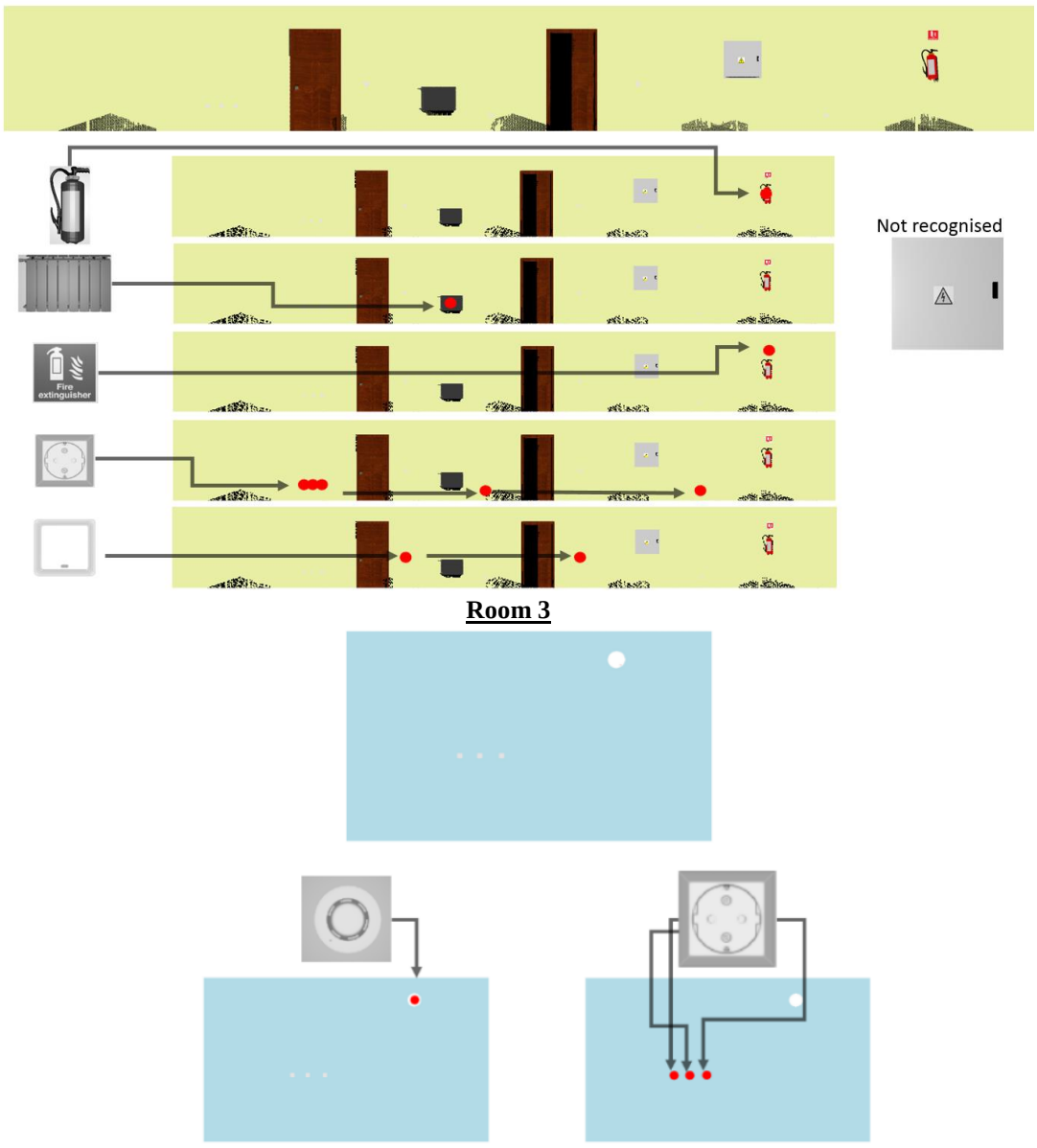
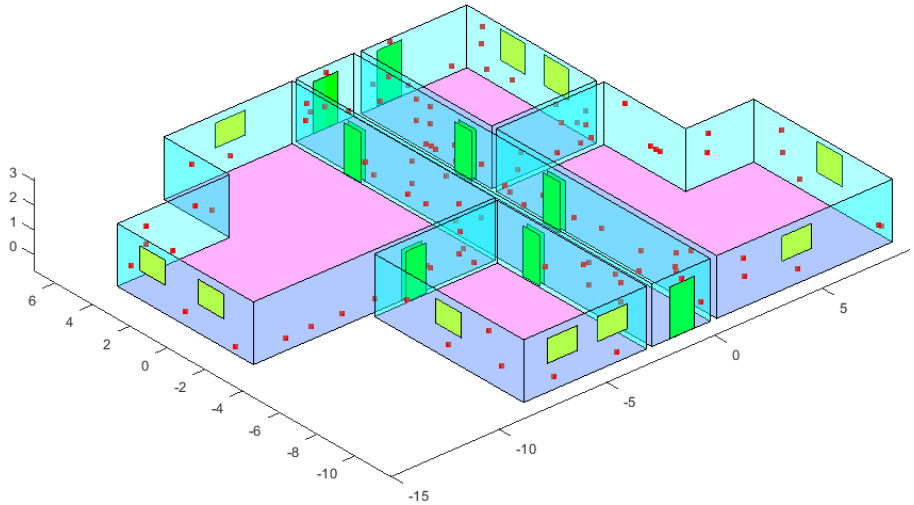
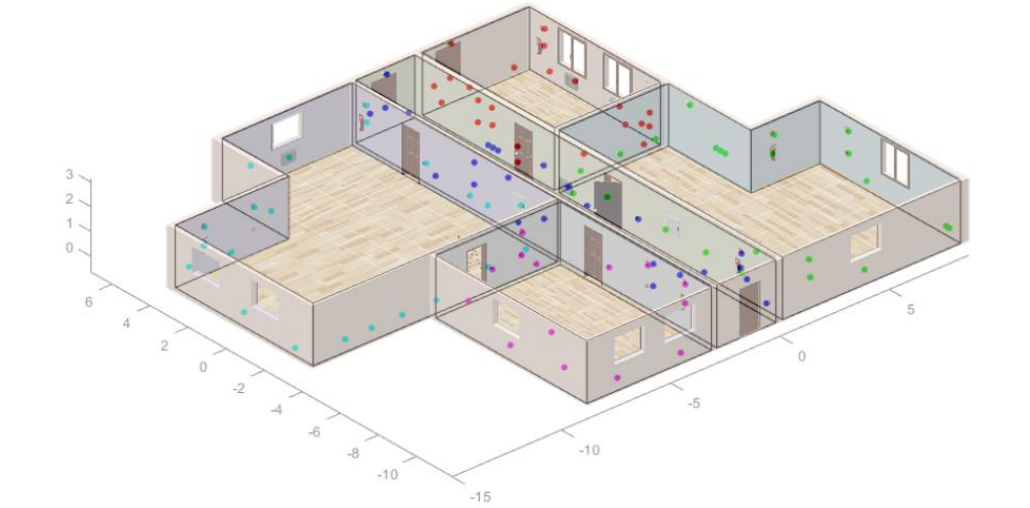


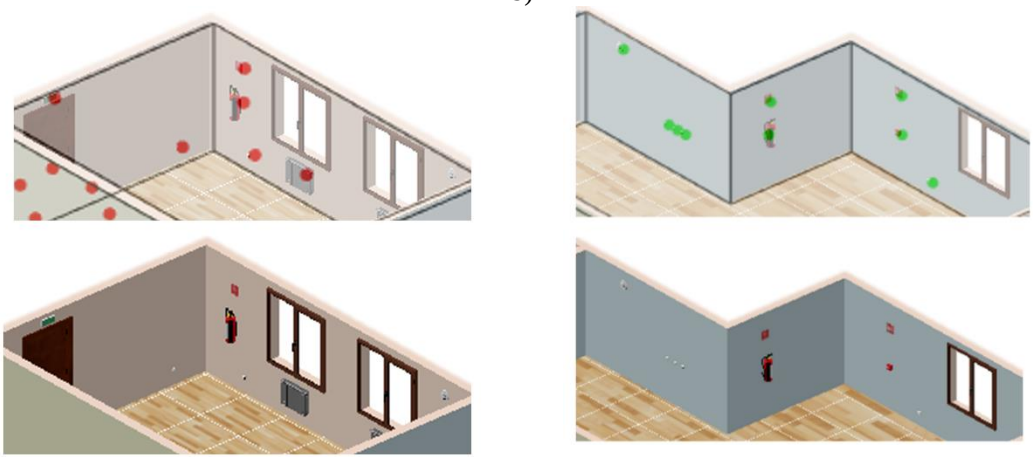
Figure 18. Examples of the recognized objects on three representative walls. Red spots indicate the position of the recognized object together with the corresponding object model.



a)



b)



c)

Figure 19. Recognition results for the simulated scenario. a) Structural elements (walls, doors and openings) and the recognized secondary objects (red spots). Doors and windows are previously recognized. b) Recognition results superimposed onto the 3D textured model in Blender. The spots are painted here in different colours depending on the room. c) Details of the earlier 3D model for a better visualization.

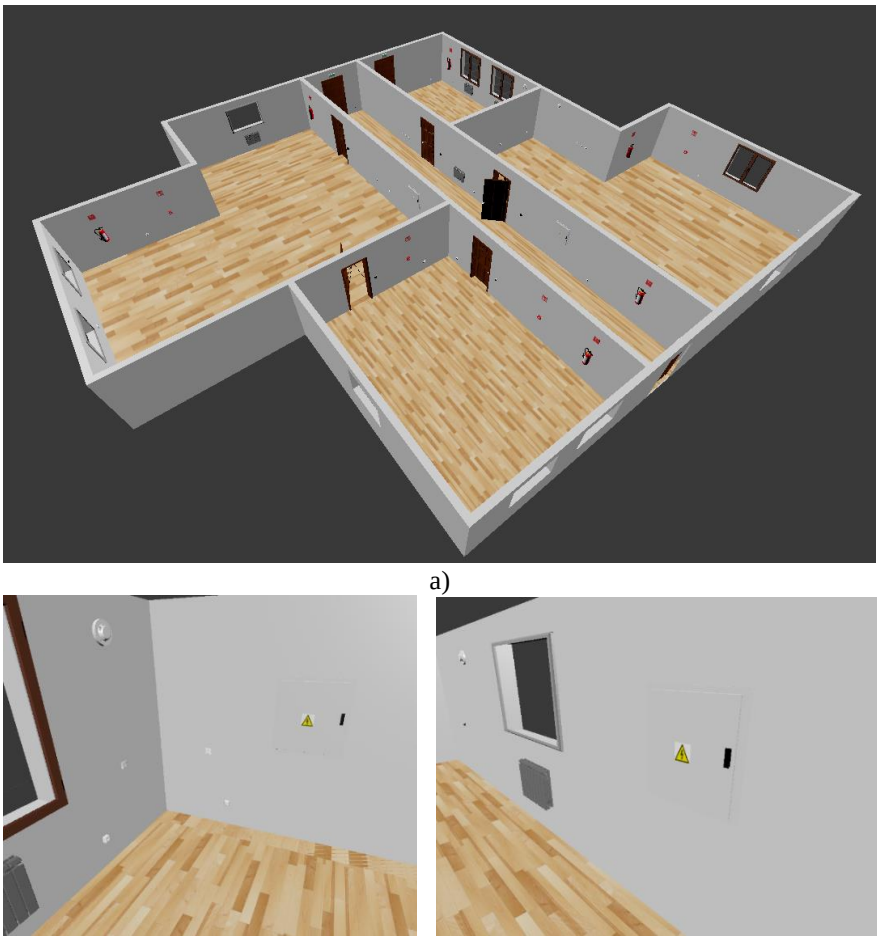
Scenario II. Minimizing the SBSC colour contrast.

In the second scenario we aim to evaluate the efficiency of our method when we force the conditions towards a minimum colour contrast between the wall and the SBSCs. The colour contrast is defined as the change in the appearance of a colour (i.e. the SBSC colour) surrounded by another colour (i.e. wall colour). In the scenario II we have drawn all the walls with a colour similar to that most of the SBSCs. The difference between both RGB components has been forced up to a minimum colour distance of 1.56% (mean distance for R, G and B components), which happens in the case of the electrical panel.

Table 2 shows the results obtained in the scenario II. A general comment is that, although we obtain worse results compared to that of the scenario I, the method maintains reasonable results. The percentage of true positives has decreased from 90.5% to 80.2%, whereas the false positives and false negatives have raised to 19.8% and 15.5% respectively. In contrast, we observe no variation on the horizontal and vertical mean errors, which remain below 3 mm..

It is clear from the values of the sixth column (TP with $\alpha>0.9$) that some of the SBSCs are not recognized because of the low colour contrast. Thus, the recognition percentages of all sockets have slightly decreased (seven points on average) with respect that of the scenario I. However, the percentages corresponding to the electrical panel and the build-in socket have

fallen to zero. This signifies that, owing to the low colour contrast between the object and the wall, both objects are not recognized by the colour-based recognition algorithm.



a)

b)

Figure 20. Scenario II. a) 3D view of the scenario. b) Details of some walls with SBSCs. Note the low colour contrast between some objects and the wall. Particularly, the contour of the electrical panel makes hardly distinguishable from the wall.

Table 2. Scenario II. Recognition results for building service components

Object	Number of instances	TP (%)	FN (%)	FP (%)	TP (%) α>0.9	$\overline{\Delta_h}$ (mm)	$\overline{\Delta_v}$ (mm)	Max Δ_h (mm)	Max Δ_v (mm)
Electrical Panel	4	50.0	50.0	0.0	0.0	2.5	3.0	5	3
Socket x1	20	70.0	30.0	15.0	28.6	1.4	2.5	3	6
Socket x2	6	100.0	0.0	0.0	83.3	2.2	3.0	5	5
Socket x4	11	100.0	0.0	0.0	54.5	2.0	2.4	5	4
Built-in Socket	6	83.3	16.7	16.7	0.0	1.8	2.2	3	5
Switch	16	87.5	12.5	12.5	100.0	2.4	2.2	4	6
Fire Extinguisher	9	100.0	0.0	0.0	44.4	13.6	7.1	48	35
Radiator	6	100.0	0.0	0.0	0.0	2.8	1.8	5	5
Fire Alarm Switch	8	62.5	37.5	37.5	60.0	1.2	3.2	4	4
Smoke Detector	10	70.0	30.0	30.0	85.7	1.6	1.6	4	3
Exit Light	3	100.0	0.0	0.0	0.0	0.0	14.7	0	23
Extinguisher Sign	9	77.8	22.2	22.2	0.0	0.9	0.9	3	3
Fire Alarm Switch Sign	8	50.0	50.0	50.0	0.0	2.8	2.3	5	3
TOTAL/Average	116	80.2	19.8	15.5	45.2	2.8	2.9	48	35

5.2. Real Scenario

Our approach has been tested on real walls using coloured point clouds collected from a Riegl VZ400 laser scanner associated with a DLSR camera on a robotic platform. This is our MoPAD (Mobile Platform for Autonomous Digitization) platform. In this section, we analyse the results obtained in a representative case of study. The orthoimage J_{CD} from Figure 1 corresponds to one of the walls of the Industrial Engineering School at Castilla La Mancha University. J_{CD} was obtained after scanning the corridor from two different locations. This is a good and complex case study that contains several objects, some of which can only be detected by colour.

The object depth models were built in our lab by using a Minolta 910 laser scanner and a turntable, while the colour models were extracted from public websites. As can be seen in Figure 1, the wall contains the following objects: one extinguisher, one fire-alarm switch, one alarm sign, one extinguisher sign and an exit sign.

Table 3 presents the recognition and positioning results of the test. The values of parameters ρ_D and ρ_C of the objects recognized after the consensus stage are in the third and fourth columns. The highest values of α are achieved in two cases (Extinguisher 2 and fire-alarm switch), in which the object is correctly recognized in both $\hat{J}_D$and $\hat{J}_C$. Items #2 and #5 were exclusively recognized using colour (signs), due to the lack of depth discontinuities for these models. On the contrary item #4 was incorrectly recognized. In this case, the RoIs corresponding to the extinguisher sign and fire alarm switch sign have quite similar feature patterns and the colour-based recognition algorithm failed. In summary, the overall recognition rate was 80%.

With regard to the localization accuracy, mean errors $\overline{\Delta_h}$ and $\overline{\Delta_v}$ are below 10 mm, which can be considered an accurate positioning result. Figure 21 a) shows the RoIs in depth (cyan) and colour (magenta), whereas Figure 21 b) illustrates the results obtained after the consensus procedure.

Table 3. Results of the experimental test for the real scenario

Item #	Object	ρ_D	ρ_C	α	$\overline{\Delta_h}$ (mm)	$\overline{\Delta_v}$ (mm)	Correctly Recognized
1	Fire alarm Switch	0,80	0,80	1,00	5	10	Yes
2	Exit Sign	-	0,90	0,50	2	2	Yes
3	Extinguisher 2	0,82	0,84	1,00	0	3	Yes
4	Extinguisher Sign 1	-	0,89	0,50	-	-	No
5	Fire Alarm Switch Sign	-	0,86	0,50	1	1	Yes

a)

b)

Figure 21. a) RoIs eventually assigned to potential objects. Depth RoIs are in cyan) and colour RoIs are in magenta. b) Results obtained after the consensus approach. The coloured spots represent the centroid of the recognized object. Note that the extinguisher sign was recognized as fire alarm switch sign.

6. CONCLUSIONS

The automatic generation of 3D BIM models is a cutting-edge research line as regards 3D computer vision when applied to buildings. In the last few years, various approaches for the recognition of structural elements of buildings (i.e. walls, ceilings, floors and openings) have been proposed. Nevertheless, the automatic recognition and positioning of small and secondary building service components in BIM models is a challenging issue on which very little research has been carried out to date.

This paper presents a new 6D-based (3D coordinates + RGB) approach that processes dense coloured 3D points with the aim of recognizing small components in buildings. This fusion of imaged and geometric algorithms is a new strategy in the automatic creation of semantically-

rich 3D models, (i.e. BIM models). Our approach makes several contributions to the field of the automatic BIM modelling.

We have developed an automatic method which does not require any human intervention throughout the entire process. Moreover, whereas most of the approaches recognize, at most, among three/four objects, we deal with a higher number of different objects that are usually in buildings.

The use of colour and depth information, along with a consensus strategy stage, makes our method more effective and robust than others. This is because some small components that are not detectable in colour images (for example, white sockets installed on white walls) can be recognized in the geometric space, and vice versa. Our algorithm is, therefore, able to recognize objects such as sockets, switches, signs, alarm devices, extinguishers and others on walls. This is an original idea that has never been seen before in the semantic 3D modelling field.

Additionally, beyond the usual recognition and pose in 2D images or partial point clouds, our method calculates the precise position of the recognized object in a 3D scenario, so that the reconstructed as-is 3D BIM model of the building can be augmented with this information about secondary components.

Nonetheless, our method has some limitations that should be dealt with in the future. Some improvements are outlined below.

The training of the system for the colour recognition stage has been carried out by defining global descriptors that are invariant to scale and rotation, which is not the case in the recognition carried out with the depth information algorithm. It will, therefore, be necessary to redefine this algorithm in order to improve the geometric recognition in cases in which the objects are rotated in the orthoimage.

The output of our system is the position (i.e. 3D coordinates) of the objects recognized in the 3D BIM model, but no mention is made of the non-recognized object. The absence of these objects could be owing to errors in the building construction phase and it is possible that the system does not recognize them because they are not in fact inside the building. In a similar way, non-expected objects in the as-designed BIM model could be correctly detected by our approach if we were to extend our model database. All these exciting issues require a response in our future work.

Finally, with regard to the experimental tests, we have presented experimentation in a simulated building and on a representative example of real wall. Although we have tested the algorithm on many other isolated real walls, with other types of secondary objects, we aim to carry out an experimentation on an entire building and so achieve, like in Figure 19, complete semantic 3D models of real buildings. In the future, our recognition and positioning algorithm should therefore be tested in more complex real scenes (e. g. curved walls and objects in columns) and complete buildings, with several instances per object on walls and in more extensive object databases.

7. ACKNOWLEDGMENTS

This work has been supported by the Spanish Economy and Competitiveness Ministry [DPI2016-76380-R project], by Castilla-La Mancha Government [PEII-2014-017-P project] and by the University of Castilla-La Mancha [PREDUCLM16/23 human resources grant].

8. REFERENCES

[1] buildingSMART, “Investors Report: Building Information Modelling (BIM),” London,

2010.

[2] "Industrial strategy: government and industry in partnership," 2012.

[3] C. M. Eastman, *BIM handbook : a guide to building information modeling for owners, managers, designers, engineers and contractors*. Wiley, 2011.

[4] E. A. Pärn, D. J. Edwards, and M. C. P. Sing, "The building information modelling trajectory in facilities management: A review," *Autom. Constr.*, vol. 75, pp. 45–55, Mar. 2017.

[5] P. Pishdad-Bozorgi, X. Gao, C. Eastman, and A. P. Self, "Planning and developing facility management-enabled building information model (FM-enabled BIM)," *Autom. Constr.*, vol. 87, pp. 22–38, Mar. 2018.

[6] S. M. Z. Borgsen, M. Schöpfer, L. Ziegler, and S. Wachsmuth, "Automated door detection with a 3D-sensor," *Proc. - Conf. Comput. Robot Vision, CRV 2014*, pp. 276–282, 2014.

[7] K. M. Varadarajan and M. Vincze, "3D room modeling and doorway detection from indoor stereo imagery using feature guided piecewise depth diffusion," *IEEE/RSJ 2010 Int. Conf. Intell. Robot. Syst. IROS 2010 - Conf. Proc.*, pp. 2758–2765, 2010.

[8] T. H. Yuan, F. H. Hashim, W. M. D. W. Zaki, and A. B. Huddin, "An Automated 3D Scanning Algorithm using Depth Cameras for Door Detection," *2015 Int. Electron. Symp.*, pp. 58–61, 2015.

[9] E. Valero, A. Adan, D. Huber, and C. Cerrada, "Detection, Modeling, and Classification of Moldings for Automated Reverse Engineering of Buildings from 3D Data," in *International Symposium on Automation and Robotics in Construction (ISARC)*, 2011.

[10] T. Czerniawski, M. Nahangi, C. Haas, and S. Walbridge, "Pipe spool recognition in cluttered point clouds using a curvature-based shape descriptor," *Autom. Constr.*, vol. 71, pp. 346–358, Nov. 2016.

[11] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 1996, pp. 226–231.

[12] P. Kim, J. Chen, Y. K. Cho, P. Kim, J. Chen, and Y. Kwon, "Building element recognition with thermal- mapped point clouds Building element recognition with thermal-mapped point clouds," in *34th International Symposium on Automation and Robotics in Construction (ISARC 2017)*, 2017, no. June.

[13] P. Kim, B. Jingdao Chen, and B. K. Yong Cho, "Robotic sensing and object recognition from thermal-mapped point clouds," *Int. J. Intell. Robot. Appl.*, vol. 1, 2017.

[14] L. Díaz-Vilariño, H. González-Jorge, J. Martínez-Sánchez, and H. Lorenzo, "Automatic LiDAR-based lighting inventory in buildings," *Meas. J. Int. Meas. Confed.*, vol. 73, pp. 544–550, 2015.

[15] I. Puente, H. González-Jorge, J. Martínez-Sánchez, and P. Arias, "Automatic detection of road tunnel luminaires using a mobile LiDAR system," *Meas. J. Int. Meas. Confed.*, vol. 47, no. 1, pp. 569–575, 2014.

[16] V. Eruhimov and W. Meeussen, "Outlet detection and pose estimation for robot continuous operation," *IEEE Int. Conf. Intell. Robot. Syst.*, pp. 2941–2946, 2011.

[17] W. Meeussen *et al.*, "Autonomous door opening and plugging in with a personal robot," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 729–736, 2010.

[18] U. Krispel, H. L. Evers, M. Tamke, R. Viehauser, and D. W. Fellner, “Automatic texture and orthophoto generation from registered panoramic views,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 40, no. 5W4, pp. 131–137, 2015.

[19] J. G. Kang, S. Y. An, W. S. Choi, and S. Y. Oh, “Recognition and path planning strategy for autonomous navigation in the elevator environment,” *Int. J. Control. Autom. Syst.*, vol. 8, no. 4, pp. 808–821, 2010.

[20] H. Hamledari, B. McCabe, and S. Davari, “Automated computer vision-based detection of components of under-construction indoor partitions,” *Autom. Constr.*, vol. 74, pp. 78–94, 2017.

[21] T. M. Bonanni, A. Pennisi, D. Bloisi, L. Iocchi, and D. Nardi, “Human-Robot Collaboration for Semantic Labeling of the Environment,” in *Proceedings of the 3rd Workshop on Semantic Perception, Mapping and Exploration*, 2013, no. July 2015, pp. 1–25.

[22] T. Czerniawski, M. Nahangi, C. Haas, and S. Walbridge, “Pipe spool recognition in cluttered point clouds using a curvature-based shape descriptor,” *Autom. Constr.*, vol. 71, pp. 346–358, Nov. 2016.

[23] H. Hamledari, B. Mccabe, and S. Davari, “Automated computer vision-based detection of components of under-construction indoor partitions,” 2017.

[24] B. Quintana, S. A. Prieto, A. Adán, and A. S. Vázquez, “Semantic Scan Planning for Indoor Structural Elements of Buildings,” *Adv. Eng. Informatics*, 2016.

[25] B. Quintana, S. A. Prieto, A. Adán, and F. Bosché, “Door detection in 3D coloured point clouds of indoor environments,” *Autom. Constr.*, vol. 85, pp. 146–166, 2018.

[26] M. Gschwandtner, R. Kwitt, A. Uhl, and W. Pree, “BlenSor: Blender sensor simulation toolbox,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6939 LNCS, no. PART 2, pp. 199–208, 2011.